122982

Schreiber, David

From: Davis, Minh-Tam
Sent: Thursday, May 20, 2004 12:48 PM
To: Schreiber, David
Subject: FW: Rush search request for 10/048046

A more complete request, with Room number and phone.
Sorry.
Tam

7/1999

-----Original Message-----
From: Davis, Minh-Tam
Sent: Thursday, May 20, 2004 12:46 PM
To: Schreiber, David
Subject: Rush search request for 10/048046

Please perform a ~~Score/leng~~th sequence for complement of SEQ ID NO:1, residues that encode amino acids 31-103, or amino acids 303-346, or amino acids 476-641 of SEQ ID NO:2.
No specific minimum hit length is required. Maximum hit length 50
Score ~~over length value of 60% or greater.~~
Thank you.
MINH TAM DAVIS
ART UNIT 1642,
ROOM 3A24, MB 3C18
272-0830

100

303-346 997-1128
476-641 1516-2013

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 20, 2004, 13:18:38 ; Search time 2 Seconds
(without alignments)
1.779 Million cell updates/sec

Title: us-10-048-046-2
Perfect score: 3585
Sequence: 1 MERPEEGKQSPPPQPWGRLL.....VKAHHAMKFNHICEQTRFKN 664

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1 seqs, 2679 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=us-10-048-046-2 -DB=us-10-048-046-1
-SUFFIX=pto -OUT=align2_1 -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits
-START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : us-10-048-046-1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3585	100.0	2679	1	us-10-048-046-1	

ALIGNMENTS

RESULT 1
us-10-048-046-1

Alignment Scores:

Pred. No.:	0	Length:	2679
Score:	3585.00	Matches:	664
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

us-10-048-046-2 (1-664) x us-10-048-046-1 (1-2679)

```
Qy          1 MetGluArgProGluGluGlyLysGlnSerProProProGlnProTrpGlyArgLeuLeu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ATGGAGCGGCCCGAGGAAGGCAAGCAGTCGCCGCCGCCGCAGCCCTGGGGACGGCTCCTG 150

Qy         21 ArgLeuGlyAlaGluGluGlyGluProHisValLeuLeuArgLysArgGluTrpThrIle 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CGTCTGGGCGCGGAGGAGGGCGAGCCGCACGTCCTCCTGAGGAAGCGGGAGTGGACCATC 210

Qy         41 GlyArgArgArgGlyCysAspLeuSerPheProSerAsnLysLeuValSerGlyAspHis 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GGGCGGAGACGAGGTTGCGACCTTTCCTTCCCCAGCAATAAACTGGTCTCTGGAGATCAC 270

Qy         61 CysArgIleValValAspGluLysSerGlyGlnValThrLeuGluAspThrSerThrSer 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 TGTAGAATTGTAGTGGATGAAAAATCAGGTCAGGTGACACTGGAAGATACCAGCACCAGT 330

Qy         81 GlyThrValIleAsnLysLeuLysValValLysLysGlnThrCysProLeuGlnThrGly 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GGAACAGTGATTAACAAGCTGAAGGTTGTTAAGAAGCAGACATGCCCTTTACAGACTGGG 390

Qy        101 AspValIleTyrLeuValTyrArgLysAsnGluProGluHisAsnValAlaTyrLeuTyr 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GATGTCATCTACTTGGTGTACAGGAAGAATGAACCGGAACACAACGTGGCATACCTCTAT 450

Qy        121 GluSerLeuSerGluLysGlnGlyMetThrGlnGluSerPheGluAlaAsnLysGluAsn 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 GAATCTTTAAGTGAAAAGCAAGGCATGACACAAGAATCCTTTGAAGCTAACAAGGAAAAT 510

Qy        141 ValPheHisGlyThrLysAspThrSerGlyAlaGlyAlaGlyArgGlyAlaAspProArg 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GTGTTCCATGGGACCAAAGATACCTCAGGTGCAGGTGCAGGGCGAGGGGCCGATCCCCGG 570

Qy        161 ValProProSerSerProAlaThrGlnValCysPheGluGluProGlnProSerThrSer 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 GTCCCTCCGTCGTCGCCCGCCACTCAGGTGTGCTTTGAGGAACCACAGCCATCAACATCG 630

Qy        181 ThrSerAspLeuProThrAlaSerAlaSerSerThrGluProSerProAlaGlyArg 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 ACGTCAGACCTCTTCCCCACAGCCTCGGCCTCTTCCACGGAGCCTTCTCCTGCAGGGCGA 690

Qy        201 GluArgSerSerSerCysGlySerGlyGlyGlyGlyIleSerProLysGlySerGlyPro 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 GAGCGTTCCTCCAGTTGTGGGTCTGGGGGTGGTGGCATCTCCCCTAAAGGAAGTGGTCCC 750

Qy        221 SerValAlaSerAspGluValSerSerPheAlaSerAlaLeuProAspArgLysThrAla 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 TCTGTGGCAAGTGATGAAGTCTCCAGCTTTGCCTCAGCTCTCCCAGACAGAAAGACTGCG 810

Qy        241 SerPheSerSerLeuGluProGlnAspGlnGluAspLeuGluProValLysLysLysMet 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 TCCTTTTCGTCGTTGGAACCCCAGGATCAGGAGGATTTGGAGCCCGTGAAGAAGAAAATG 870

Qy        261 ArgGlyAspGlyAspLeuAspLeuAsnGlyGlnLeuLeuValAlaGlnProArgArgAsn 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 AGAGGAGATGGGGACCTTGACCTGAACGGGCAGTTGTTGGTCGCACAACCGCGTAGAAAT 930

Qy        281 AlaGlnThrValHisGluAspValArgAlaAlaAlaGlyLysProAspLysMetGluGlu 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 GCCCAAACCGTCCACGAGGACGTCAGAGCAGCGGCTGGGAAGCCAGACAAGATGGAGGAG 990

Qy        301 ThrLeuThrCysIleIleCysGlnAspLeuLeuHisAspCysValSerLeuGlnProCys 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        991 ACGCTGACATGCATCATCTGCCAGGACCTGCTGCACGACTGCGTGAGTTTGCAGCCCTGC 1050

Qy        321 MetHisThrPheCysAlaAlaCysTyrSerGlyTrpMetGluArgSerSerLeuCysPro 340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 ATGCACACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCCT 1110

Qy        341 ThrCysArgCysProValGluArgIleCysLysAsnHisIleLeuAsnAsnLeuValGlu 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1111 ACCTGCCGCTGTCCCGTGGAGCGGATCTGTAAAAACCACATCCTCAACAACCTCGTGGAA 1170

Qy        361 AlaTyrLeuIleGlnHisProAspLysSerArgSerGluGluAspValGlnSerMetAsp 380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1171 GCATACCTCATCCAGCATCCAGACAAGAGTCGCAGTGAAGAAGATGTGCAAAGTATGGAT 1230
```

```
Qy        381 AlaArgAsnLysIleThrGlnAspMetLeuGlnProLysValArgArgSerPheSerAsp 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1231 GCCAGGAATAAAATCACTCAAGACATGCTGCAGCCCAAAGTCAGGCGGTCTTTTTCTGAT 1290

Qy        401 GluGluGlySerSerGluAspLeuLeuGluLeuSerAspValAspSerGluSerSerAsp 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 GAAGAAGGGAGTTCAGAGGACCTGCTGGAGCTGTCAGACGTTGACAGTGAGTCCTCAGAC 1350

Qy        421 IleSerGlnProTyrValValCysArgGlnCysProGluTyrArgArgGlnAlaAlaGln 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1351 ATTAGCCAGCCATACGTCGTGTGCCGGCAGTGTCCTGAGTACAGAAGGCAGGCGGCGCAG 1410

Qy        441 ProProHisCysProAlaProGluGlyGluProGlyAlaProGlnAlaLeuGlyAspAla 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 CCTCCCCACTGCCCAGCACCCGAGGGCGAGCCAGGAGCCCCACAGGCCCTGGGGGATGCA 1470

Qy        461 ProSerThrSerValSerLeuThrThrAlaValGlnAspTyrValCysProLeuGlnGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 CCCTCCACGTCCGTCAGCCTGACGACAGCAGTCCAGGATTACGTGTGCCCTCTGCAAGGA 1530

Qy        481 SerHisAlaLeuCysThrCysCysPheGlnProMetProAspArgArgAlaGluArgGlu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 AGCCACGCCCTGTGCACCTGCTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGCGAG 1590

Qy        501 GlnAspProArgValAlaProGlnGlnCysAlaValCysLeuGlnProPheCysHisLeu 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 CAGGACCCGCGTGTCGCCCCTCAGCAGTGTGCGGTCTGCCTGCAGCCTTTCTGCCACCTG 1650

Qy        521 TyrTrpGlyCysThrArgThrGlyCysTyrGlyCysLeuAlaProPheCysGluLeuAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 TACTGGGGCTGCACCCGGACCGGCTGCTACGGCTGCCTGGCCCCGTTTTGTGAGCTCAAC 1710

Qy        541 LeuGlyAspLysCysLeuAspGlyValLeuAsnAsnAsnSerTyrGluSerAspIleLeu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1711 CTGGGTGACAAGTGTCTGGACGGCGTGCTGAACAACAACAGCTACGAGTCAGACATCCTG 1770

Qy        561 LysAsnTyrLeuAlaThrArgGlyLeuThrTrpLysAsnMetLeuThrGluSerLeuVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 AAGAATTACCTGGCAACCAGAGGTTTGACATGGAAAAACATGTTGACCGAGAGCCTCGTG 1830

Qy        581 AlaLeuGlnArgGlyValPheLeuLeuSerAspTyrArgValThrGlyAspThrValLeu 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 GCTCTCCAGCGGGGAGTGTTTCTGCTGTCTGATTACAGAGTCACGGGAGACACCGTTCTG 1890

Qy        601 CysTyrCysCysGlyLeuArgSerPheArgGluLeuThrTyrGlnTyrArgGlnAsnIle 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 TGTTACTGCTGTGGCCTGCGCAGCTTCCGTGAGCTGACCTATCAGTATCGGCAGAACATT 1950

Qy        621 ProAlaSerGluLeuProValAlaValThrSerArgProAspCysTyrTrpGlyArgAsn 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 CCTGCTTCCGAGTTGCCAGTGGCCGTAACATCCCGTCCTGACTGCTACTGGGGCCGTAAC 2010

Qy        641 CysArgThrGlnValLysAlaHisHisAlaMetLysPheAsnHisIleCysGluGlnThr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 TGCCGCACTCAGGTGAAAGCTCACCACGCCATGAAATTCAATCATATCTGTGAACAGACA 2070

Qy        661 ArgPheLysAsn 664
              ||||||||||||
Db       2071 AGGTTCAAAAAC 2082
```

Search completed: May 20, 2004, 13:18:43
Job time : 5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:10:28 ; Search time 3707.14 Seconds
(without alignments)
5822.506 Million cell updates/sec

Title: US-10-048-046-1_COPY_1516_2013
Perfect score: 498
Sequence: 1 tgccctctgcaaggaagcca.....gctactggggccgtaactgc 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	4.2	30	6	I13395	I13395 Sequence 5
	2	21	4.2	30	6	I13775	I13775 Sequence 5
	3	21	4.2	30	6	I14876	I14876 Sequence 3
	4	21	4.2	30	6	I68745	I68745 Sequence 5
	5	21	4.2	30	6	I92686	I92686 Sequence 3
	6	21	4.2	30	6	AR437164	AR437164 Sequence
	7	20.8	4.2	50	6	AX165879	AX165879 Sequence
	8	20.4	4.1	47	6	BD169427	BD169427 Genomes p
	9	20.2	4.1	50	6	AR403372	AR403372 Sequence
	10	20.2	4.1	50	6	BD068872	BD068872 Enzymatic
c	11	19.8	4.0	40	6	AX456252	AX456252 Sequence
	12	19.8	4.0	42	6	AR078407	AR078407 Sequence
	13	19.8	4.0	50	6	AR403397	AR403397 Sequence
	14	19.8	4.0	50	6	BD068897	BD068897 Enzymatic
c	15	19.6	3.9	48	6	AR277861	AR277861 Sequence
c	16	19.4	3.9	42	6	E54530	E54530 Herpesvirus
	17	19	3.8	36	6	AR206156	AR206156 Sequence
c	18	19	3.8	42	6	AR116992	AR116992 Sequence
	19	19	3.8	42	6	AR116993	AR116993 Sequence
	20	19	3.8	44	9	HUMSPARC01	M25737 Human osteo
	21	18.8	3.8	50	6	I18452	I18452 Sequence 75
c	22	18.6	3.7	36	9	S81380	S81380 T cell anti
	23	18.6	3.7	40	6	BD180755	BD180755 Array of
	24	18.6	3.7	42	6	AX411814	AX411814 Sequence
c	25	18.6	3.7	47	6	AX004825	AX004825 Sequence
c	26	18.6	3.7	47	6	BD074255	BD074255 Method fo
	27	18.4	3.7	36	6	BD271168	BD271168 Virus vec
c	28	18.4	3.7	36	6	BD271169	BD271169 Virus vec
	29	18.4	3.7	36	6	AR264600	AR264600 Sequence
c	30	18.4	3.7	36	6	AR264601	AR264601 Sequence
c	31	18.4	3.7	38	6	AR057811	AR057811 Sequence
c	32	18.4	3.7	38	6	AR057818	AR057818 Sequence
c	33	18.4	3.7	38	6	AR057975	AR057975 Sequence
c	34	18.4	3.7	38	6	AR115569	AR115569 Sequence
c	35	18.4	3.7	38	6	AR115576	AR115576 Sequence
c	36	18.4	3.7	38	6	AR115733	AR115733 Sequence
c	37	18.4	3.7	38	6	AX634870	AX634870 Sequence
c	38	18.4	3.7	38	6	AX634877	AX634877 Sequence
c	39	18.4	3.7	38	6	AX635034	AX635034 Sequence
c	40	18.4	3.7	41	6	AX518482	AX518482 Sequence
	41	18.4	3.7	45	6	AR282656	AR282656 Sequence
c	42	18.4	3.7	45	6	AR282657	AR282657 Sequence
	43	18.4	3.7	45	6	AX370679	AX370679 Sequence
c	44	18.4	3.7	45	6	AX370680	AX370680 Sequence
	45	18.4	3.7	47	6	AR289352	AR289352 Sequence

ALIGNMENTS

RESULT 1
I13395
LOCUS I13395 30 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 5 from patent US 5436153.
ACCESSION I13395
VERSION I13395.1 GI:910736
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Sprecher,C.A., Foster,D.C. and Norris,K.E.
TITLE Human amyloid protein precursor homolog and Kunitz-type inhibitor
JOURNAL Patent: US 5436153-A 5 25-JUL-1995;
FEATURES Location/Qualifiers

```
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match              4.2%;  Score 21;  DB 6;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 1.5e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 2
I13775
LOCUS       I13775                    30 bp    DNA     linear   PAT 26-SEP-1995
DEFINITION  Sequence 5 from patent US 5441931.
ACCESSION   I13775
VERSION     I13775.1  GI:996205
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Sprecher,C.A., Foster,D.C. and Norris,K.E.
  TITLE     Human amyloid protein precursor homologue and Kunitz-type
            inhibitors
  JOURNAL   Patent: US 5441931-A 5 15-AUG-1995;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match              4.2%;  Score 21;  DB 6;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 1.5e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 3
I14876
LOCUS       I14876                    30 bp    DNA     linear   PAT 02-APR-1996
DEFINITION  Sequence 3 from patent US 5455338.
ACCESSION   I14876
VERSION     I14876.1  GI:1249785
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Sprecher,C.A., Kisiel,W. and Foster,D.C.
  TITLE     DNA encoding novel human kunitz-type inhibitors and methods
            relating thereto
  JOURNAL   Patent: US 5455338-A 3 03-OCT-1995;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match              4.2%;  Score 21;  DB 6;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 1.5e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 4
I68745
LOCUS       I68745                    30 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 5 from patent US 5677146.
ACCESSION   I68745
VERSION     I68745.1  GI:2830867
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Sprecher,C.A., Foster,D.C. and Norris,K.E.
  TITLE     Human amyloid protein precursor homolog and kunitz-type inhibitor
  JOURNAL   Patent: US 5677146-A 5 14-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match              4.2%;  Score 21;  DB 6;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 1.5e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 5
I92686
LOCUS       I92686                    30 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 3 from patent US 5728674.
ACCESSION   I92686
VERSION     I92686.1  GI:3937156
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Sprecher,C.A., Kisiel,W. and Foster,D.C.
  TITLE     Inhibition of blood coagulation by human-kunitz-type inhibitors
  JOURNAL   Patent: US 5728674-A 3 17-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match              4.2%;  Score 21;  DB 6;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 1.5e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 6
AR437164
LOCUS       AR437164                  30 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 3 from patent US 6656746.
ACCESSION   AR437164
VERSION     AR437164.1  GI:40200261
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Sprecher,C.A., Kisiel,W. and Foster,D.C.
  TITLE     Antibodies to human kunitz-type inhibitor
  JOURNAL   Patent: US 6656746-A 3 02-DEC-2003;
```

```
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match              4.2%;  Score 21;  DB 6;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 1.5e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              ||| |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 7
AX165879
LOCUS       AX165879                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1074 from Patent WO0138586.
ACCESSION   AX165879
VERSION     AX165879.1  GI:14546708
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 1074 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43250166"
     variation       26
                     /note="single nucleotide polymorphism"
ORIGIN

  Query Match              4.2%;  Score 20.8;  DB 6;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 1.7e+06;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        454 GTGGCCGTAACATCCCGTCCTGACTGCTACTGGGGCCGTA 493
              |||||| | ||| || || ||| ||||     |||| |||
Db          3 GTGGCCCTCACACCCAGTGCTGTCTGCGCGGAGGGCTGTA 42


RESULT 8
BD169427
LOCUS       BD169427                 47 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Genomes participating in rheumatoid arthritis, method of diagnosing
            the same, method of judging the onset risk thereof, kit for
            detecting and diagnosing the same, method of treating rheumatoid
            arthritis and remedies therefor.
ACCESSION   BD169427
VERSION     BD169427.1  GI:27875239
KEYWORDS    WO 0234912-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Shiozawa,S. and Konishi,Y.
  TITLE     Genomes participating in rheumatoid arthritis, method of diagnosing
            the same, method of judging the onset risk thereof, kit for
            detecting and diagnosing the same, method of treating rheumatoid
            arthritis and remedies therefor
  JOURNAL   Patent: WO 0234912-A 4 02-MAY-2002;
            SHUNICHI SHIOZAWA,YOSHITAKE KONISHI
COMMENT     OS   Artificial Sequence
            PN   WO 0234912-A/4
            PD   02-MAY-2002
            PF   24-OCT-2001 WO 2001JP009313
            PR   24-OCT-2000 JP 00P   324296,27-MAR-2001 JP 01P   090546 PR
            30-MAR-2001 JP 01P   099990
            PI   SHUNICHI SHIOZAWA,YOSHITAKE KONISHI
            PC   C12N15/12,C07K14/47,C12Q1/68,G01N33/50,A61K38/17,A61K48/00 CC
             Synthesized oligonucleotide
            FH   Key             Location/Qualifiers
            FT   source          1. .47
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match              4.1%;  Score 20.4;  DB 6;  Length 47;
  Best Local Similarity   65.2%;  Pred. No. 2.2e+06;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         27 GTGCACCTGCTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGC 72
              | | |||  | |||||| ||||  ||| | | ||   | |||| ||
Db          2 GGGTACCATCCGCTTCCTGCCCCAGCCAGGCTGGTTTGTGGAGTGC 47


RESULT 9
AR403372
LOCUS       AR403372                 50 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1712 from patent US 6623962.
ACCESSION   AR403372
VERSION     AR403372.1  GI:40150822
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Akhtar,S., Fell,P. and McSwiggen,J.A.
  TITLE     Enzymatic nucleic acid treatment of diseases of conditions related
            to levels of epidermal growth factor receptors
  JOURNAL   Patent: US 6623962-A 1712 23-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match              4.1%;  Score 20.2;  DB 6;  Length 50;
  Best Local Similarity   75.8%;  Pred. No. 2.4e+06;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        268 GCAACCAGAGGTTTGACATGGAAAAACATGTTG 300
              ||| ||||| ||||  ||  |||| ||| ||||
Db          1 GCACCCAGAAGTTTACCAGAGAAACACACGTTG 33


RESULT 10
BD068872
LOCUS       BD068872                 50 bp    RNA     linear   PAT 27-AUG-2002
DEFINITION  Enzymatic nucleic acid treatment of diseases or conditions related
            to levels of epidermal growth factor receptors.
ACCESSION   BD068872
VERSION     BD068872.1  GI:22614475
KEYWORDS    JP 2001511003-A/1712.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Akhtar,S., Fell,P. and Mcswiggen,J.A.
  TITLE     Enzymatic nucleic acid treatment of diseases or conditions related
```

```
            to levels of epidermal growth factor receptors
  JOURNAL   Patent: JP 2001511003-A 1712 07-AUG-2001;
            RIBOZYME PHARMACEUTICALS INC,ASTON UNIV
COMMENT     OS   Unidentified
            PN   JP 2001511003-A/1712
            PD   07-AUG-2001
            PF   14-JAN-1998 JP 1998532913
            PR   31-JAN-1997 US    60/036476,04-DEC-1997 US    08/985162 PI
            SAGHIR AKHTAR,PATRICIA FELL,JAMES A MCSWIGGEN PC
            C12N9/00,C07K14/71
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Enzymatic nucleic acid treatment of diseases or conditions  CC
                 related to
            CC   levels of epidermal growth factor receptors
            FH   Key             Location/Qualifiers
            FT   source          1. .50
            FT                   /organism='Unidentified'.
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="unidentified"
                      /mol_type="genomic RNA"
                      /db_xref="taxon:32644"
ORIGIN

  Query Match              4.1%;  Score 20.2;  DB 6;  Length 50;
  Best Local Similarity   75.8%;  Pred. No. 2.4e+06;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        268 GCAACCAGAGGTTTGACATGGAAAAACATGTTG 300
              ||| ||||| ||||  ||  |||| ||| ||||
Db          1 GCACCCAGAAGTTTACCAGAGAAACACACGTTG 33


RESULT 11
AX456252/c
LOCUS       AX456252                 40 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 110 from Patent WO0216944.
ACCESSION   AX456252
VERSION     AX456252.1  GI:21715166
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Wood,K.V., Wood,M.G., Zhuang,Y. and Paguio,A.
  TITLE     Synthetic nucleic acid molecule compositions and methods of
            preparation
  JOURNAL   Patent: WO 0216944-A 110 28-FEB-2002;
            PROMEGA CORPORATION (US)
FEATURES              Location/Qualifiers
     source           1. .40
                      /organism="synthetic construct"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32630"
                      /note="An oligonucleotide"
ORIGIN

  Query Match              4.0%;  Score 19.8;  DB 6;  Length 40;
  Best Local Similarity   69.2%;  Pred. No. 3e+06;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        372 TCTGTGTTACTGCTGTGGCCTGCGCAGCTTCCGTGAGCT 410
              | ||  | ||   | ||  ||| |||||| ||||||||
Db         39 TTTGGTTGACAAGTATGATCTGAGCAGCTTGCGTGAGCT 1


RESULT 12
AR078407
LOCUS       AR078407                 42 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 26 from patent US 5962636.
ACCESSION   AR078407
VERSION     AR078407.1  GI:10005153
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Bachmaier,K., Hessel,A.John., Neu,N. and Penninger,J.Martin.
  TITLE     Peptides capable of modulating inflammatory heart disease
  JOURNAL   Patent: US 5962636-A 26 05-OCT-1999;
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="unknown"
                      /mol_type="unassigned DNA"
ORIGIN

  Query Match              4.0%;  Score 19.8;  DB 6;  Length 42;
  Best Local Similarity   77.4%;  Pred. No. 3e+06;
  Matches   24;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        115 CTGCAGCCTTTCTGCCACCTGTACTGGGGCT 145
              | ||| ||| | ||||| || | ||||||||
Db         12 CAGCACCCTGTTTGCCAACTATGCTGGGGCT 42


RESULT 13
AR403397
LOCUS       AR403397                 50 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1737 from patent US 6623962.
ACCESSION   AR403397
VERSION     AR403397.1  GI:40150847
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Akhtar,S., Fell,P. and McSwiggen,J.A.
  TITLE     Enzymatic nucleic acid treatment of diseases of conditions related
            to levels of epidermal growth factor receptors
  JOURNAL   Patent: US 6623962-A 1737 23-SEP-2003;
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="unknown"
                      /mol_type="genomic DNA"
ORIGIN

  Query Match              4.0%;  Score 19.8;  DB 6;  Length 50;
  Best Local Similarity   63.8%;  Pred. No. 3e+06;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy        224 TGAACAACAACAGCTACGAGTCAGACATCCTGAAGAATTACCTGGCA 270
              |||| || |||| |   ||  || || |  ||     ||||||||||
Db          4 TGAAGAAGAACACCAGAGAAACACACGTTGTGGTACATTACCTGGTA 50


RESULT 14
BD068897
LOCUS       BD068897                 50 bp    RNA     linear   PAT 27-AUG-2002
DEFINITION  Enzymatic nucleic acid treatment of diseases or conditions related
            to levels of epidermal growth factor receptors.
ACCESSION   BD068897
VERSION     BD068897.1  GI:22614500
KEYWORDS    JP 2001511003-A/1737.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Akhtar,S., Fell,P. and Mcswiggen,J.A.
  TITLE     Enzymatic nucleic acid treatment of diseases or conditions related
            to levels of epidermal growth factor receptors
  JOURNAL   Patent: JP 2001511003-A 1737 07-AUG-2001;
            RIBOZYME PHARMACEUTICALS INC,ASTON UNIV
COMMENT     OS   Unidentified
            PN   JP 2001511003-A/1737
            PD   07-AUG-2001
```

```
            PF   14-JAN-1998 JP 1998532913
            PR   31-JAN-1997 US    60/036476,04-DEC-1997 US    08/985162 PI
            SAGHIR AKHTAR,PATRICIA FELL,JAMES A MCSWIGGEN PC
            C12N9/00,C07K14/71
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Enzymatic nucleic acid treatment of diseases or conditions  CC
                 related to
            CC   levels of epidermal growth factor receptors
            FH   Key             Location/Qualifiers
            FT   source          1. .50
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unidentified"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             4.0%;  Score 19.8;  DB 6;  Length 50;
  Best Local Similarity   63.8%;  Pred. No. 3e+06;
  Matches   30;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy        224 TGAACAACAACAGCTACGAGTCAGACATCCTGAAGAATTACCTGGCA 270
              |||| || |||| |   ||  || || |  ||    ||||||||| |
Db          4 TGAAGAAGAACACCAGAGAAACACACGTTGTGGTACATTACCTGGTA 50


RESULT 15
AR277861/c
LOCUS       AR277861                 48 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 20 from patent US 6511803.
ACCESSION   AR277861
VERSION     AR277861.1  GI:29711785
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Church,G.M. and Mitra,R.D.
  TITLE     Replica amplification of nucleic acid arrays
  JOURNAL   Patent: US 6511803-A 20 28-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             3.9%;  Score 19.6;  DB 6;  Length 48;
  Best Local Similarity   73.5%;  Pred. No. 3.4e+06;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         36 CTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAG 69
              || |||| |||    ||||||  ||| |||||||
Db         43 CTCCTTCAAGCTTCCGCCCGAGAGGAAAGCGGAG 10


Search completed: May 26, 2004, 07:39:20
Job time : 3709.14 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:01:38 ; Search time 420.572 Seconds
(without alignments)
5030.293 Million cell updates/sec

Title: US-10-048-046-1_COPY_1516_2013
Perfect score: 498
Sequence: 1 tgcctctgcaaggaagcca.....gctactggggccgtaactgc 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	21.6	4.3	31	2	AAX06374	Aax06374 Human bia
c	2	21.2	4.3	50	2	AAV04790	Aav04790 Primer 95
c	3	21.2	4.3	50	2	AAV31045	Aav31045 Expressio
	4	21	4.2	30	2	AAQ66249	Aaq66249 Sequence
	5	21	4.2	30	2	AAQ90470	Aaq90470 Human Kun
	6	21	4.2	50	4	AAL32458	Aal32458 Human SNP
	7	21	4.2	50	4	AAL32043	Aal32043 Human SNP
c	8	20.8	4.2	41	6	ABV74248	Abv74248 Human mac
	9	20.8	4.2	48	7	ACC62681	Acc62681 Human ant
	10	20.8	4.2	48	7	ACC62643	Acc62643 Human ant
	11	20.8	4.2	50	5	ABL01083	Abl01083 Human SNP
	12	20.4	4.1	47	6	AAL47188	Aal47188 Human rhe
	13	20.2	4.1	50	2	AAV98932	Aav98932 Human EGF
c	14	20	4.0	50	6	ABZ00921	Abz00921 Human leu
c	15	19.8	4.0	40	6	ABL99142	Abl99142 Green/red
	16	19.8	4.0	42	2	AAZ28188	Aaz28188 Human bet
	17	19.8	4.0	42	3	AAZ99171	Aaz99171 Human pep
	18	19.8	4.0	50	2	AAV98957	Aav98957 Human EGF
c	19	19.8	4.0	50	4	AAL34581	Aal34581 Human SNP
	20	19.8	4.0	50	4	AAL34023	Aal34023 Human SNP
c	21	19.6	3.9	41	6	ABA05129	Aba05129 Human syn
c	22	19.6	3.9	48	3	AAA99461	Aaa99461 Primer Pu
c	23	19.6	3.9	48	4	AAS12549	Aas12549 PCR prime
	24	19.6	3.9	50	6	ABZ07831	Abz07831 Human leu
c	25	19.2	3.9	41	6	ABV74249	Abv74249 Human mac
	26	19.2	3.9	50	4	AAL28785	Aal28785 Human SNP
c	27	19.2	3.9	50	4	AAL34331	Aal34331 Human SNP
c	28	19.2	3.9	50	4	AAL34262	Aal34262 Human SNP
	29	19	3.8	50	4	AAL34327	Aal34327 Human SNP
	30	18.8	3.8	41	6	ABL55510	Abl55510 Human pal
	31	18.8	3.8	48	7	ACC62644	Acc62644 Human ant
	32	18.8	3.8	48	7	ACC62693	Acc62693 Human ant
	33	18.8	3.8	48	7	ACC62655	Acc62655 Human ant
	34	18.8	3.8	48	7	ACC62682	Acc62682 Human ant
	35	18.6	3.7	42	6	AAD39247	Aad39247 pD10-CMV-
c	36	18.6	3.7	44	2	AAQ51167	Aaq51167 DNA fragm
c	37	18.6	3.7	47	2	AAX56654	Aax56654 Primer 27
	38	18.6	3.7	50	6	ABX03168	Abx03168 HCV hairp
	39	18.6	3.7	50	6	ABX03263	Abx03263 HCV hairp
	40	18.6	3.7	50	6	ABZ02873	Abz02873 Human leu
	41	18.4	3.7	36	3	AAD00854	Aad00854 Primer #1
c	42	18.4	3.7	36	3	AAD00855	Aad00855 Primer #2
c	43	18.4	3.7	38	2	AAT53816	Aat53816 Rat ICAM
c	44	18.4	3.7	38	2	AAT53973	Aat53973 Rat ICAM
c	45	18.4	3.7	38	2	AAT53809	Aat53809 Rat ICAM

ALIGNMENTS

RESULT 1
AAX06374/c
ID AAX06374 standard; DNA; 31 BP.
XX
AC AAX06374;
XX
DT 31-MAR-1999 (first entry)
XX
DE Human biallelic polymorphic DNA fragment SGC33319.
XX
KW Polymorphism; biallelic; paternity testing; forensic; genetic mapping;
KW phenotypic typing; medicament; disease; marker; human; ss.
XX
OS Homo sapiens.
XX
PN WO9858529-A2.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-US012930.
XX
PR 24-JUN-1997; 97US-0050594P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Lipshutz RJ, Chee M, Fan J, Berno A;
XX
DR WPI; 1999-080963/07.
XX
PT New nucleic acid segments containing polymorphic sites - used for, e.g.
PT detecting a disease phenotype, in forensics, paternity testing or genetic
PT mapping of phenotypic traits.
XX
PS Claim 1; Page 23; 61pp; English.
XX
CC Sequences AAX06101-X06558 represent human DNA fragments which contain
CC biallelic polymorphic markers. The base occupying the polymorphic site is
CC indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
CC can be used in a method for determining polymorphic forms in an
CC individual. The invention further provides computer-readable storage
CC medium for storing data for access by an application programme being
CC executed on a data processing system. Such a method comprises a data
CC structure stored in the computer-readable storage medium, the data
CC structure including information resident in a database used by the
CC application programme and including records, each record comprising
CC information identifying a polymorphism shown in the above sequences. The

CC products and methods can be used for analysing polymorphic sites in
CC individuals for testing for the presence of a disease phenotype or in
CC forensics, paternity testing or genetic mapping of phenotypic traits.
CC They can also be used for the production of polypeptides expressed by
CC variant genes and for the production of transgenic animals. The nucleic
CC acid segments can also be used in the manufacture of medicaments for the
CC treatment or prophylaxis of diseases
XX
SQ Sequence 31 BP; 5 A; 8 C; 8 G; 9 T; 0 U; 1 Other;

Query Match 4.3%; Score 21.6; DB 2; Length 31;
Best Local Similarity 80.0%; Pred. No. 2.5e+04;
Matches 24; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 230 ACAACAGCTACGAGTCAGACATCCTGAAGA 259
||| |||||||||||:| | |||||| |||
Db 30 ACATCAGCTACGAGKCTGGAATCCTGGAGA 1

RESULT 2
AAV04790/c
ID AAV04790 standard; cDNA; 50 BP.
XX
AC AAV04790;
XX
DT 10-JUN-1998 (first entry)
XX
DE Primer 95-1414 for green fluorescent protein.
XX
KW PCR primer; green fluorescent protein; GFP; recombinant fungal host cell;
KW cryptic splice site; hormone production; enzyme production;
KW receptor production; reporter production; ss.
XX
OS Synthetic.
OS Aequorea victoria.
XX
PN WO9749821-A1.
XX
PD 31-DEC-1997.
XX
PF 20-JUN-1997; 97WO-US010630.
XX
PR 27-JUN-1996; 96US-0020930P.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Thompson SA;
XX
DR WPI; 1998-077184/07.
XX
PT Recombinant heterologous polypeptide expression in fungal host cells -
PT comprises modifying at least 1 cryptic splice site in nucleic acid
PT sequence encoding heterologous polypeptide, to provide correct
PT expression.
XX
PS Example 1; Page 22; 59pp; English.
XX
CC This sequence represents a primer used in the construction of a vector
CC containing the Aequorea victoria green fluorescent protein (GFP). The
CC vector can be used in the method of the invention for obtaining a novel
CC recombinant fungal host cell (FHC), which comprises introducing a nucleic
CC acid sequence (NAS) encoding a heterologous polypeptide, where at least 1
CC cryptic splice site (CSS) is modified, into a FHC. The products can be
CC used for the recombinant production of hormones, enzymes (e.g.
CC aminopeptidases, amylases, carbohydrases, carboxypeptidases, catalases,
CC cellulases, chitinases, cutinases, deoxyribonucleases, esterases, alpha-
CC galactosidases, beta-galactosidases, glucoamylases, alpha-glucosidases,
CC beta-glucosidases, haloperoxidases, invertases, laccases, lipases,
CC mannosidases, mutanases, oxidases, pectinolytic enzymes, peroxidases,
CC phytases, polyphenoloxidases, proteolytic enzymes, ribonucleases or
CC xylanases), receptors or reporters (e.g. Aequorea victoria GFP). The CSS
CC modification prevents improper splicing of precursor mRNA for
CC heterologous expression by FHC, therefore preventing the production of
CC biologically inactive polypeptides, or the production of several
CC populations of polypeptide products with varying biological activity
XX
SQ Sequence 50 BP; 8 A; 13 C; 18 G; 11 T; 0 U; 0 Other;

Query Match 4.3%; Score 21.2; DB 2; Length 50;
Best Local Similarity 69.0%; Pred. No. 3.8e+04;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 396 CAGCTTCCGTGAGCTGACCTATCAGTATCGGCAGAACATTCC 437
| || | | |||| || | || |||| |||||||| |||
Db 50 CGGCAGCGTTCAGCTCGCCGACCATTATCAGCAGAACACTCC 9

RESULT 3
AAV31045/c
ID AAV31045 standard; cDNA; 50 BP.
XX
AC AAV31045;
XX
DT 17-AUG-1998 (first entry)
XX
DE Expression vector pShTh49 construction primer 95-1414.
XX
KW Fungal host cell; cryptic splice site; filamentous fungus; primer;
KW splicing; biologically active; construction; expression vector; ss.
XX
OS Synthetic.
OS Aequorea victoria.
XX
PN US5770371-A.
XX
PD 23-JUN-1998.
XX
PF 27-JUN-1996; 96US-00672158.
XX
PR 27-JUN-1996; 96US-00672158.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Thompson S;
XX
DR WPI; 1998-376791/32.
XX
PT DNA for transforming filamentous fungi - with cryptic splice sites
PT modified to prevent incorrect splicing in fungi.
XX
PS Example 1; Col 14; 27pp; English.
XX
CC The present sequence represents a primer used in the construction of the
CC expression vector pShTh49 in an example from the present invention. The
CC present invention describes a nucleic acid sequence with at least one
CC modified cryptic splice site, where the nucleic acid sequence encodes a
CC biologically active polypeptide and the cryptic splice site is modified
CC by (a) replacing at least one cryptic consensus sequence with a non-
CC consensus sequence; and/or (b) replacing a first region of a cryptic
CC intron with a second region that has a G+C content of 40-70%. The fungal
CC host cells may be used to produce biologically active polypeptides,
CC preferably hormones, enzymes, receptors or reporters. Problems of
CC incorrect splicing of precursor mRNA in filamentous fungal cells can be
CC overcome
XX
SQ Sequence 50 BP; 8 A; 13 C; 18 G; 11 T; 0 U; 0 Other;

Query Match 4.3%; Score 21.2; DB 2; Length 50;
Best Local Similarity 69.0%; Pred. No. 3.8e+04;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 396 CAGCTTCCGTGAGCTGACCTATCAGTATCGGCAGAACATTCC 437
| || | | |||| || | || |||| |||||||| |||
Db 50 CGGCAGCGTTCAGCTCGCCGACCATTATCAGCAGAACACTCC 9

```
RESULT 4
AAQ66249
ID   AAQ66249 standard; cDNA; 30 BP.
XX
AC   AAQ66249;
XX
DT   25-MAR-2003  (revised)
DT   29-DEC-1994  (first entry)
XX
DE   Sequence of probe ZC4792 for screening for a protease inhibitor from the
DE   Kunitz family from human placenta.
XX
KW   Kunitz-type protease inhibitor; serine protease inhibitor; therapy;
KW   amyloid precursor; probe; ss.
XX
OS   Synthetic.
XX
PN   WO9412637-A2.
XX
PD   09-JUN-1994.
XX
PF   01-DEC-1993;   93WO-US011696.
XX
PR   02-DEC-1992;   92US-00985692.
PR   19-NOV-1993;   93US-00155331.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Sprecher CA,  Foster DC,  Norris K;
XX
DR   WPI; 1994-200265/24.
XX
PT   New human Kunitz type protease inhibitor and related DNA - for treating
PT   pancreatitis and other disorders involving serine protease, also new
PT   amyloid protein precursor homologues including the inhibitor in its
PT   sequence.
XX
PS   Example; Page 47; 70pp; English.
XX
CC   A DNA segment encoding a human Kunitz-type inhibitor which comprises bps
CC   171-331 or 159-331 of AAQ66247 is claimed. Also claimed is the DNA
CC   encoding AAR53776 from AA 56-110, from 53-110 or from 57-110 and which
CC   further contains a Glu on the NH2 terminus. Isolated human Kunitz- type
CC   inhibitor which comprises the aforementioned sequences is also claimed.
CC   To obtain a cDNA sequence encoding a protease inhibitor from the Kunitz
CC   family, a human placenta cDNA library in lambda gt11 was screened using
CC   the radio-labelled probe ZC4792 (AAQ66249). Plasmids contg. two resultant
CC   clones - ZGKI13 and ZGKI20 - were deposited on October 14th 1992 with the
CC   ATCC under accession nos. 69090 and 69089 respectively. Plasmids ZGKI13
CC   and ZGKI20 were shown to contain the identical partial sequence shown in
CC   AAQ66247 and AAR53776. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 30 BP; 2 A; 10 C; 8 G; 10 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 21;  DB 2;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 3.7e+04;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 5
AAQ90470
ID   AAQ90470 standard; DNA; 30 BP.
XX
AC   AAQ90470;
XX
DT   21-JAN-1996  (first entry)
XX
DE   Human Kunitz-type protease-inhibitor antisense DNA probe ZC4792.
XX
KW   Kunitz-type inhibitor; protease-inhibitor; enzyme-inhibitor;
KW   anticoagulant; deep vein thrombosis; antisense DNA probe; ss.
XX
OS   Synthetic.
XX
PN   WO9512674-A1.
XX
PD   11-MAY-1995.
XX
PF   02-NOV-1994;   94WO-US012609.
XX
PR   05-NOV-1993;   93US-00147710.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
PA   (UYNE-) UNIV NEW MEXICO STATE.
XX
PI   Sprecher CA,  Kisiel W,  Foster DC;
XX
DR   WPI; 1995-193821/25.
XX
PT   Human Kunitz-type protease inhibitors - used as anticoagulants and in the
PT   treatment of deep vein thrombosis.
XX
PS   Disclosure; Page 51; 65pp; English.
XX
CC   Poly(A)+ RNAs from a variety of human tissues were screened using this
CC   antisense 30-mer oligonucleotide as an initial step in the cloning of a
CC   novel human Kunitz inhibitor cDNA
XX
SQ   Sequence 30 BP; 2 A; 10 C; 8 G; 10 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 21;  DB 2;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 3.7e+04;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 6
AAL32458
ID   AAL32458 standard; DNA; 50 BP.
XX
AC   AAL32458;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #5666.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US035498.
XX
PR   28-DEC-1999;   99US-0173419P.
PR   27-DEC-2000; 2000US-00173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
```

PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Claim 1; Page 3019; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-
CC protein coupled receptors and thioesterases. The present sequence is one
CC such oligonucleotide. The oligonucleotides and the peptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of the proteins listed above.
CC Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms
XX
SQ Sequence 50 BP; 12 A; 17 C; 10 G; 11 T; 0 U; 0 Other;

Query Match 4.2%; Score 21; DB 4; Length 50;
Best Local Similarity 73.0%; Pred. No. 4.3e+04;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy        125 TCTGCCACCTGTACTGGGGCTGCACCCGGACCGGCTG 161
              ||||||||||  ||  | || ||| |||   |||||
Db         13 TCTGCCACCTGAGCTAAGCCTACACACGGCGTGGCTG 49
```

RESULT 7
AAL32043
ID AAL32043 standard; DNA; 50 BP.
XX
AC AAL32043;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5251.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Claim 1; Page 2898; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-
CC protein coupled receptors and thioesterases. The present sequence is one
CC such oligonucleotide. The oligonucleotides and the peptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of the proteins listed above.
CC Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms
XX
SQ Sequence 50 BP; 8 A; 15 C; 17 G; 10 T; 0 U; 0 Other;

Query Match 4.2%; Score 21; DB 4; Length 50;
Best Local Similarity 82.8%; Pred. No. 4.3e+04;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         24 CCTGTGCACCTGCTGCTTCCAGCCCATGC 52
              | || |||||||||||||| ||||  |||
Db         11 CATGGGCACCTGCTGCTTCTAGCCACTGC 39
```

RESULT 8
ABV74248/c
ID ABV74248 standard; DNA; 41 BP.
XX
AC ABV74248;
XX
DT 02-JAN-2003 (first entry)
XX
DE Human macro protein 0986-9.24 probe SEQ ID NO 8.
XX
KW Human; macro protein 0986-9.24; tumour; haemopathy; development disorder;
KW human immunodeficiency virus; HIV; infection immunological disease;
KW inflammation; probe; ss.
XX
OS Homo sapiens.
XX
PN CN1352149-A.
XX
PD 05-JUN-2002.
XX
PF 10-NOV-2000; 2000CN-00127337.
XX
PR 10-NOV-2000; 2000CN-00127337.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-699451/76.
XX
PT New polypeptide-human macro protein 0986-9.24 and polynucleotide for
PT encoding such polypeptide.
XX
PS Example 6; Page 19 (Disclosure); 32pp; Chinese.
XX
CC The invention relates to human macro protein 0986-9.24, polynucleotides
CC encoding this polypeptide and DNA recombinant processes to produce the
CC polypeptide. The present invention also discloses applying the

```
CC   polypeptide in treating various diseases, such as malignant tumours,
CC   haemopathy, development disorder, human immunodeficiency virus (HIV)
CC   infection, immunological diseases and various inflammations. The present
CC   sequence is that of a human macro protein 0986-9.24 probe, useful in
CC   examples of the invention
XX
SQ   Sequence 41 BP; 8 A; 9 C; 13 G; 11 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 20.8;  DB 6;  Length 41;
  Best Local Similarity   70.0%;  Pred. No. 4.6e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        420 GTATCGGCAGAACATTCCTGCTTCCGAGTTGCCAGTGGCC 459
              | || |||| ||||||||  |    ||| | |||||| ||
Db         41 GCATGGGCAAAACATTCCATCCCTAGAGGTTCCAGTGCCC 2


RESULT 9
ACC62681
ID   ACC62681 standard; DNA; 48 BP.
XX
AC   ACC62681;
XX
DT   30-JUN-2003  (first entry)
XX
DE   Human antibody heavy chain framework 1 primer TMX24VH1b SEQ ID NO:64.
XX
KW   Engineered template; single primer amplification; antibody library;
KW   nucleic acid amplification; primer; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO2003025202-A2.
XX
PD   27-MAR-2003.
XX
PF   19-SEP-2002; 2002WO-US029889.
XX
PR   19-SEP-2001; 2001US-0323455P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS,  Frederickson S,  Maruyama T,  Lin Y,  Renshaw M;
XX
DR   WPI; 2003-313359/30.
XX
PT   Amplifying nucleic acid by contacting engineered nucleic acid strand
PT   having predetermined sequence at one end and sequence complementary to
PT   predetermined sequence at other end, with primer having predetermined
PT   sequence.
XX
PS   Example 3; Page 29; 68pp; English.
XX
CC   The present invention describes a method (M1) for amplifying a nucleic
CC   acid strand. M1 comprises providing an engineered nucleic acid strand (S)
CC   having a predetermined sequence at one end and a sequence complementary
CC   to the predetermined sequence at the other, and contacting (S) with a
CC   primer having the predetermined sequence in the presence of a polymers
CC   and nucleotides under conditions suitable for polymerisation of the
CC   nucleotides. Also described is an engineered nucleic acid strand (I)
CC   having a predetermined sequence at one end and a sequence complementary
CC   to the predetermined sequence at the other end. M1 is useful for
CC   amplifying a nucleic acid. M1 can be used for producing an antibody
CC   library. M1 is useful for preparing amplified products that can be
CC   ligated into a suitable expression vector, where the vector can be used
CC   to transform an appropriate host organism to produce the polypeptide or
CC   protein encoded by the target sequence. M1 is useful to amplify a family
CC   of related sequences to build a complex library such as, for example an
CC   antibody library. M1 is useful not only for producing large amounts of
CC   one target nucleic acid sequence, but also for amplifying simultaneously
CC   more than one different target nucleic acid sequence located on the same
CC   or different nucleic acid molecules. ACC62635 to ACC62753 and ABR54841 to
CC   ABR54998 represent sequence used in the exemplification of the present
CC   invention
XX
SQ   Sequence 48 BP; 8 A; 10 C; 20 G; 10 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 20.8;  DB 7;  Length 48;
  Best Local Similarity   70.0%;  Pred. No. 4.8e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || |||||||| || || |||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGCAGGTCCAGCTTGT 44


RESULT 10
ACC62643
ID   ACC62643 standard; DNA; 48 BP.
XX
AC   ACC62643;
XX
DT   30-JUN-2003  (first entry)
XX
DE   Human antibody heavy chain VH framework 1 primer TMX24VH1b SEQ ID NO:26.
XX
KW   Engineered template; single primer amplification; antibody library;
KW   nucleic acid amplification; primer; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO2003025202-A2.
XX
PD   27-MAR-2003.
XX
PF   19-SEP-2002; 2002WO-US029889.
XX
PR   19-SEP-2001; 2001US-0323455P.
XX
PA   (ALEX-) ALEXION PHARM INC.
XX
PI   Bowdish KS,  Frederickson S,  Maruyama T,  Lin Y,  Renshaw M;
XX
DR   WPI; 2003-313359/30.
XX
PT   Amplifying nucleic acid by contacting engineered nucleic acid strand
PT   having predetermined sequence at one end and sequence complementary to
PT   predetermined sequence at other end, with primer having predetermined
PT   sequence.
XX
PS   Example 2; Page 23; 68pp; English.
XX
CC   The present invention describes a method (M1) for amplifying a nucleic
CC   acid strand. M1 comprises providing an engineered nucleic acid strand (S)
CC   having a predetermined sequence at one end and a sequence complementary
CC   to the predetermined sequence at the other, and contacting (S) with a
CC   primer having the predetermined sequence in the presence of a polymers
CC   and nucleotides under conditions suitable for polymerisation of the
CC   nucleotides. Also described is an engineered nucleic acid strand (I)
CC   having a predetermined sequence at one end and a sequence complementary
CC   to the predetermined sequence at the other end. M1 is useful for
CC   amplifying a nucleic acid. M1 can be used for producing an antibody
CC   library. M1 is useful for preparing amplified products that can be
CC   ligated into a suitable expression vector, where the vector can be used
CC   to transform an appropriate host organism to produce the polypeptide or
CC   protein encoded by the target sequence. M1 is useful to amplify a family
CC   of related sequences to build a complex library such as, for example an
CC   antibody library. M1 is useful not only for producing large amounts of
CC   one target nucleic acid sequence, but also for amplifying simultaneously
CC   more than one different target nucleic acid sequence located on the same
CC   or different nucleic acid molecules. ACC62635 to ACC62753 and ABR54841 to
CC   ABR54998 represent sequence used in the exemplification of the present
CC   invention
XX
```

```
SQ   Sequence 48 BP; 8 A; 10 C; 20 G; 10 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 20.8;  DB 7;  Length 48;
  Best Local Similarity   70.0%;  Pred. No. 4.8e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || |||||||| || || |||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGCAGGTCCAGCTTGT 44


RESULT 11
ABL01083
ID   ABL01083 standard; DNA; 50 BP.
XX
AC   ABL01083;
XX
DT   05-MAR-2002  (first entry)
XX
DE   Human SNP involving a gap oligonucleotide SEQ ID NO:1074.
XX
KW   Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW   immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW   autoimmune disease; inflammation; cancer; nervous system disease;
KW   infection; polymorphic protein; ds.
XX
OS   Homo sapiens.
XX
PN   WO200138586-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032311.
XX
PR   24-NOV-1999;   99US-0167383P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-355949/37.
XX
PT   Isolated human nucleic acids comprising one or more single nucleotide
PT   polymorphisms, useful for treating a subject suffering from a pathology,
PT   e.g. autoimmune diseases, ascribed to the presence of a sequence
PT   polymorphism.
XX
PS   Claim 1; Page 575; 674pp; English.
XX
CC   ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC   comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC   to ABB56903 represent human peptides encoded by some of the SNP
CC   oligonucleotides. The sequences from the present invention can have
CC   immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC   antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC   and antibodies from the present invention can be used for treating a
CC   subject suffering from, at risk for, or suspected of, suffering from a
CC   pathology ascribed to the presence of a sequence polymorphism. The
CC   pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC   the nervous system, and infection by pathogenic microorganisms. The SNPs
CC   are also useful for determining which forms of a characterised
CC   polymorphism are present in individuals. The antibodies may be used in
CC   the detection, quantitation and/or cellular or tissue localisation of a
CC   polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC   protein within appropriate physiological samples)
XX
SQ   Sequence 50 BP; 8 A; 14 C; 18 G; 10 T; 0 U; 0 Other;

  Query Match             4.2%;  Score 20.8;  DB 5;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 4.9e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        454 GTGGCCGTAACATCCCGTCCTGACTGCTACTGGGGCCGTA 493
              |||||| | ||| || || ||| ||||     |||| |||
Db          3 GTGGCCCTCACACCCAGTGCTGTCTGCGCGGAGGGCTGTA 42


RESULT 12
AAL47188
ID   AAL47188 standard; DNA; 47 BP.
XX
AC   AAL47188;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human rheumatoid arthritis associated DR3 gene related oligo #1.
XX
KW   Human; DR3; rheumatoid arthritis; antirheumatic; antiarthritic;
KW   gene therapy; PCR; primer; ss.
XX
OS   Unidentified.
XX
PN   WO200234912-A1.
XX
PD   02-MAY-2002.
XX
PF   24-OCT-2001; 2001WO-JP009313.
XX
PR   24-OCT-2000; 2000JP-00324296.
PR   27-MAR-2001; 2001JP-00090546.
PR   30-MAR-2001; 2001JP-00099990.
XX
PA   (NEWI-) NEW IND RES ORG.
PA   (SHIO/) SHIOZAWA S.
XX
PI   Shiozawa S,  Konishi Y;
XX
DR   WPI; 2002-417132/44.
XX
PT   Genomes, particularly DR3 genomic DNA, participating in rheumatoid
PT   arthritis via mutation, useful in evaluating disease onset and its
PT   possibility and providing therapy and remedies.
XX
PS   Example 4; Page 25; 84pp; Japanese.
XX
CC   The present invention relates to the human DR3 gene, which is associated
CC   with rheumatoid arthritis. Certain mutations in the gene can be linked to
CC   the disease. The sequences can be used to evaluate disease onset and its
CC   possibility and to provide therapy and remedies. The present sequence is
CC   an oligonucleotide described in the exemplification of the invention
XX
SQ   Sequence 47 BP; 5 A; 15 C; 16 G; 11 T; 0 U; 0 Other;

  Query Match             4.1%;  Score 20.4;  DB 6;  Length 47;
  Best Local Similarity   65.2%;  Pred. No. 6.1e+04;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         27 GTGCACCTGCTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGC 72
              | | |||  | |||||| ||||  ||| | | ||   | |||| ||
Db          2 GGGTACCATCCGCTTCCTGCCCCAGCCAGGCTGGTTTGTGGAGTGC 47


RESULT 13
AAV98932
ID   AAV98932 standard; RNA; 50 BP.
XX
AC   AAV98932;
XX
DT   17-MAR-1999  (first entry)
XX
DE   Human EGF-R hairpin ribozyme nucleotide position 2768.
XX
KW   Human; epidermal growth factor receptor; EGFR; EGF-R; target sequence;
KW   hammerhead ribozyme; hairpin ribozyme; inhibition; cell proliferation;
KW   cancer; genetic drift; detection; mutation; ss.
XX
```

OS Synthetic.
OS Homo sapiens.
XX
PN WO9833893-A2.
XX
PD 06-AUG-1998.
XX
PF 14-JAN-1998; 98WO-US000730.
XX
PR 31-JAN-1997; 97US-0036476P.
PR 04-DEC-1997; 97US-00985162.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (UYAS-) UNIV ASTON.
XX
PI Akhtar S, Fell P, Mcswiggen JA;
XX
DR WPI; 1998-437449/37.
XX
PT Enzymatic nucleic acids - which cleave RNA derived from an epidermal
PT growth factor receptor, useful for inhibiting cell proliferation and for
PT treating cancers.
XX
PS Claim 7; Page 88; 109pp; English.
XX
CC The present invention describes enzymatic nucleic acid molecules (NAMs)
CC which specifically cleave RNA derived from an epidermal growth factor
CC receptor (EGF-R) gene. AAV97221 to AAV98043 and AAV98979 to AAV99090
CC represent specifically claimed target sequence from human EGF-R. AAV98044
CC to AAV98866 and AAV98867 to V9878 represent hammerhead ribozymes and
CC hairpin ribozymes respectively for human EGF-R. The NAMs are useful for
CC cleaving EGF-R RNA in the treatment of a condition associated with EGFR
CC expression levels e.g. to inhibit cell proliferation in the prevention or
CC treatment of cancers. The NAMs can also be used as diagnostic tools to
CC examine genetic drift and mutations within diseased cells or to detect
CC the presence of EGF-R RNA in a cell
XX
SQ Sequence 50 BP; 16 A; 12 C; 11 G; 0 T; 11 U; 0 Other;

Query Match 4.1%; Score 20.2; DB 2; Length 50;
Best Local Similarity 60.6%; Pred. No. 7.1e+04;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 268 GCAACCAGAGGTTTGACATGGAAAAACATGTTG 300
||| ||||| |::: || |||| ||| |::|
Db 1 GCACCCAGAAGUUUACCAGAGAAACACACGUUG 33

RESULT 14
ABZ00921/c
ID ABZ00921 standard; DNA; 50 BP.
XX
AC ABZ00921;
XX
DT 09-JAN-2003 (first entry)
XX
DE Human leukocyte gene expression profiling probe SEQ ID NO 912.
XX
KW T7; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200257414-A2.
XX
PD 25-JUL-2002.
XX
PF 22-OCT-2001; 2001WO-US047856.
XX
PR 20-OCT-2000; 2000US-0241994P.
PR 08-JUN-2001; 2001US-0296764P.
XX
PA (BIOC-) BIOCARDIA INC.
XX
PI Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quertermous T, Johnson F;
XX
DR WPI; 2002-636525/68.
XX
PT New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS Claim 1; Page 353; 0pp; English.
XX
CC The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC predicting disease complications in an individual or monitoring response
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ Sequence 50 BP; 11 A; 16 C; 7 G; 16 T; 0 U; 0 Other;

Query Match 4.0%; Score 20; DB 6; Length 50;
Best Local Similarity 72.2%; Pred. No. 8.1e+04;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 198 GGGTGACAAGTGTCTGGACGGCGTGCTGAACAACAA 233
||| | |||||| |||| | | ||| ||||| ||
Db 41 GGGGCAGAAGTGTGTGGATGACCTGCCAAACAATAA 6

RESULT 15
ABL99142/c
ID ABL99142 standard; DNA; 40 BP.
XX
AC ABL99142;
XX
DT 28-JUN-2002 (first entry)
XX
DE Green/red click beetle luciferase preparing oligo SEQ ID NO:110.
XX
KW Luciferase; synthetic nucleic acid; transcriptional characteristic;
KW transcription; codon usage; PCR; primer; ss.
XX
OS Coleoptera.
OS Synthetic.
XX
PN WO200216944-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026566.
XX
PR 24-AUG-2000; 2000US-00645706.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Wood KV, Wood MG, Zhuang Y, Paguio A;
XX
DR WPI; 2002-304140/34.
XX
PT Preparing a synthetic nucleic acid molecule with reduced inappropriate
PT transcriptional characteristics when expressed in a cell, for e.g making
PT fusion proteins, by altering a wild type or another synthetic nucleic
PT acid sequence.
XX

PS Example 1; Fig 6; 294pp; English.
XX
CC The present invention relates to the preparation of synthetic nucleic
CC acid molecules which have altered transcriptional regulatory sequences
CC compared to the wild-type. These sequences are then transcribed with less
CC frequency compared to the wild-type. In particular, the invention relates
CC to altered luciferase sequences. This can be used to detect weak promoter
CC activity, to express fusion proteins, to detect and/or measure levels of
CC gene expression, subcellular localisation or targeting, in life science
CC research, agrogenetics, gene therapy, developmental science and
CC pharmaceutical development. The present sequence is an oligonucleotide
CC described in the exemplification of the invention
XX
SQ Sequence 40 BP; 13 A; 12 C; 7 G; 8 T; 0 U; 0 Other;

Query Match 4.0%; Score 19.8; DB 6; Length 40;
Best Local Similarity 69.2%; Pred. No. 8.5e+04;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy        372 TCTGTGTTACTGCTGTGGCCTGCGCAGCTTCCGTGAGCT 410
              | ||  | ||   | ||  ||| ||||||| ||||||||
Db         39 TTTGGTTGACAAGTATGATCTGAGCAGCTTGCGTGAGCT 1
```

Search completed: May 26, 2004, 05:53:55
Job time : 429.572 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:16:49 ; Search time 96.1979 Seconds
(without alignments)
2872.886 Million cell updates/sec

Title: US-10-048-046-1_COPY_1516_2013
Perfect score: 498
Sequence: 1 tgccctctgcaaggaagcca.....gctactggggccgtaactgc 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	21.2	4.3	50	1	US-08-672-158A-10	Sequence 10, Appl
	2	21	4.2	30	1	US-07-985-692-5	Sequence 5, Appli
	3	21	4.2	30	1	US-08-155-331-5	Sequence 5, Appli
	4	21	4.2	30	1	US-08-147-710-3	Sequence 3, Appli
	5	21	4.2	30	1	US-08-424-022-5	Sequence 5, Appli
	6	21	4.2	30	1	US-08-458-090-3	Sequence 3, Appli
	7	21	4.2	30	2	US-08-457-887-3	Sequence 3, Appli
	8	21	4.2	30	2	US-08-424-017B-5	Sequence 5, Appli
	9	21	4.2	30	4	US-09-904-621-3	Sequence 3, Appli
	10	21	4.2	30	5	PCT-US93-11696-5	Sequence 5, Appli
	11	20.2	4.1	50	3	US-08-985-162-1712	Sequence 1712, Ap
	12	20.2	4.1	50	4	US-09-401-063-1712	Sequence 1712, Ap
	13	19.8	4.0	42	2	US-09-133-774-26	Sequence 26, Appl
	14	19.8	4.0	42	3	US-09-303-862-26	Sequence 26, Appl
	15	19.8	4.0	50	3	US-08-985-162-1737	Sequence 1737, Ap
	16	19.8	4.0	50	4	US-09-401-063-1737	Sequence 1737, Ap
c	17	19.6	3.9	48	4	US-09-522-732B-20	Sequence 20, Appl
	18	19	3.8	36	4	US-08-327-874A-26	Sequence 26, Appl
	19	19	3.8	36	5	PCT-US94-09700-26	Sequence 26, Appl
c	20	19	3.8	42	3	US-08-448-619-4	Sequence 4, Appli
	21	19	3.8	42	3	US-08-448-619-5	Sequence 5, Appli
	22	18.8	3.8	50	1	US-07-972-032-75	Sequence 75, Appl
	23	18.4	3.7	36	4	US-09-438-268-25	Sequence 25, Appl
c	24	18.4	3.7	36	4	US-09-438-268-26	Sequence 26, Appl
c	25	18.4	3.7	38	2	US-08-292-620A-2015	Sequence 2015, Ap
c	26	18.4	3.7	38	2	US-08-292-620A-2022	Sequence 2022, Ap
c	27	18.4	3.7	38	2	US-08-292-620A-2179	Sequence 2179, Ap
c	28	18.4	3.7	38	3	US-09-071-845-2015	Sequence 2015, Ap
c	29	18.4	3.7	38	3	US-09-071-845-2022	Sequence 2022, Ap
c	30	18.4	3.7	38	3	US-09-071-845-2179	Sequence 2179, Ap
	31	18.4	3.7	45	4	US-09-765-873A-29	Sequence 29, Appl
c	32	18.4	3.7	45	4	US-09-765-873A-30	Sequence 30, Appl
	33	18.4	3.7	47	4	US-09-422-978-1087	Sequence 1087, Ap
c	34	18.4	3.7	48	4	US-09-502-558-21	Sequence 21, Appl
c	35	18.4	3.7	48	4	US-09-502-558-22	Sequence 22, Appl
c	36	18.2	3.7	46	3	US-09-372-934-14	Sequence 14, Appl
c	37	18.2	3.7	46	3	US-09-372-934-25	Sequence 25, Appl
c	38	18.2	3.7	46	4	US-09-486-241-18	Sequence 18, Appl
c	39	18.2	3.7	46	4	US-09-766-916-14	Sequence 14, Appl
c	40	18.2	3.7	46	4	US-09-766-916-25	Sequence 25, Appl
c	41	18.2	3.7	46	4	US-09-766-898-14	Sequence 14, Appl
c	42	18.2	3.7	46	4	US-09-766-898-25	Sequence 25, Appl
c	43	18.2	3.7	46	4	US-09-637-040C-14	Sequence 14, Appl
c	44	18.2	3.7	46	4	US-09-637-040C-25	Sequence 25, Appl
	45	18.2	3.7	50	3	US-08-998-099-283	Sequence 283, App

ALIGNMENTS

RESULT 1
US-08-672-158A-10/c
; Sequence 10, Application US/08672158A
; Patent No. 5770371
; GENERAL INFORMATION:
; APPLICANT: Sheryl Thompson
; TITLE OF INVENTION: Modification of Cryptic Splice Sites In
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5770371o No. 5770371disk of No. 5770371th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,158A
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 4855.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-672-158A-10

Query Match 4.3%; Score 21.2; DB 1; Length 50;
Best Local Similarity 69.0%; Pred. No. 4.9e+03;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 396 CAGCTTCCGTGAGCTGACCTATCAGTATCGGCAGAACATTCC 437
Db 50 CGGCAGCGTTCAGCTCGCCGACCATTATCAGCAGAACACTCC 9

RESULT 2

```
US-07-985-692-5
; Sequence 5, Application US/07985692
; Patent No. 5436153
;  GENERAL INFORMATION:
;    APPLICANT:  Sprecher, Cindy A
;    APPLICANT:  Foster, Donald C
;    APPLICANT:  No. 5436153ris, Kjeld
;    TITLE OF INVENTION:  NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
;    TITLE OF INVENTION:  HOMOLOG AND KUNITZ-TYPE INHIBITOR
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  4225 Roosevelt Way, N.E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/985,692
;      FILING DATE:  19921202
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-684
;      REFERENCE/DOCKET NUMBER:  92-21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-547-8080 ext 322
;      TELEFAX:  206-548-2329
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC4792
US-07-985-692-5

  Query Match             4.2%;  Score 21;  DB 1;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 3
US-08-155-331-5
; Sequence 5, Application US/08155331
; Patent No. 5441931
;  GENERAL INFORMATION:
;    APPLICANT:  Foster, Donald C
;    APPLICANT:  Sprecher, Cindy
;    APPLICANT:  No. 5441931ris, Kjeld
;    TITLE OF INVENTION:  NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
;    TITLE OF INVENTION:  HOMOLOG AND KUNITZ-TYPE INHIBITOR
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  4225 Roosevelt Way, N.E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/155,331
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,692
;      FILING DATE:  02-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-684
;      REFERENCE/DOCKET NUMBER:  92-21C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-547-8080 ext 322
;      TELEFAX:  206-548-2329
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC4792
US-08-155-331-5

  Query Match             4.2%;  Score 21;  DB 1;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 4
US-08-147-710-3
; Sequence 3, Application US/08147710
; Patent No. 5455338
;  GENERAL INFORMATION:
;    APPLICANT:  Sprecher, Cindy A.
;    APPLICANT:  Kisiel, Walter
;    APPLICANT:  Foster, Donald C.
;    TITLE OF INVENTION:  NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
;    TITLE OF INVENTION:  METHODS RELATING THERETO
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  4225 Roosevelt Way, N.E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/147,710
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-684
;      REFERENCE/DOCKET NUMBER:  93-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-547-8080 ext 322
;      TELEFAX:  206-548-2329
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       CLONE:  ZC4792
US-08-147-710-3

  Query Match             4.2%;  Score 21;  DB 1;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              ||| |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 5
US-08-424-022-5
; Sequence 5, Application US/08424022
; Patent No. 5677146
;  GENERAL INFORMATION:
;    APPLICANT:  Foster, Donald C
;    APPLICANT:  Sprecher, Cindy
;    APPLICANT:  No. 5677146ris, Kjeld
;    TITLE OF INVENTION:  NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
;    TITLE OF INVENTION:  HOMOLOG AND KUNITZ-TYPE INHIBITOR
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  4225 Roosevelt Way, N.E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,022
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,692
;      FILING DATE:  02-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-684
;      REFERENCE/DOCKET NUMBER:  92-21C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-547-8080 ext 322
;      TELEFAX:  206-548-2329
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC4792
US-08-424-022-5

  Query Match             4.2%;  Score 21;  DB 1;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              ||| |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 6
US-08-458-090-3
; Sequence 3, Application US/08458090
; Patent No. 5728674
;  GENERAL INFORMATION:
;    APPLICANT:  Sprecher, Cindy A.
;    APPLICANT:  Kisiel, Walter
;    APPLICANT:  Foster, Donald C.
;    TITLE OF INVENTION:  NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
;    TITLE OF INVENTION:  METHODS RELATING THERETO
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,090
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-648
;      REFERENCE/DOCKET NUMBER:  93-14D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6673
;      TELEFAX:  206-442-6678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC4792
US-08-458-090-3

  Query Match             4.2%;  Score 21;  DB 1;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              ||| |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 7
US-08-457-887-3
; Sequence 3, Application US/08457887
; Patent No. 5914315
;  GENERAL INFORMATION:
;    APPLICANT:  Sprecher, Cindy A.
;    APPLICANT:  Kisiel, Walter
;    APPLICANT:  Foster, Donald C.
;    TITLE OF INVENTION:  NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
;    TITLE OF INVENTION:  METHODS RELATING THERETO
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/457,887
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-648
;      REFERENCE/DOCKET NUMBER:  93-14D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6673
;      TELEFAX:  206-442-6678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC4792
US-08-457-887-3

  Query Match             4.2%;  Score 21;  DB 2;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 8
US-08-424-017B-5
; Sequence 5, Application US/08424017B
; Patent No. 5935854
;  GENERAL INFORMATION:
;    APPLICANT:  Foster, Donald C
;    APPLICANT:  Sprecher, Cindy
;    APPLICANT:  No. 5935854ris, Kjeld
;    TITLE OF INVENTION:  NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
;    TITLE OF INVENTION:  HOMOLOG AND KUNITZ-TYPE INHIBITOR
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  4225 Roosevelt Way, N.E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,017B
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/155,331
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/985,692
;      FILING DATE:  02-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-684
;      REFERENCE/DOCKET NUMBER:  92-21C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-547-8080 ext 322
;      TELEFAX:  206-548-2329
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  ZC4792
US-08-424-017B-5

  Query Match             4.2%;  Score 21;  DB 2;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 9
US-09-904-621-3
; Sequence 3, Application US/09904621
; Patent No. 6656746
; GENERAL INFORMATION:
;  APPLICANT: Sprecher, Cindy A.
;  APPLICANT:  Kisiel, Walter
;  APPLICANT:  Foster, Donald C.
;  TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS
;  TITLE OF INVENTION:  AND
;  TITLE OF INVENTION:  METHODS RELATING THERETO
;  FILE REFERENCE: 93-14D3
;  CURRENT APPLICATION NUMBER: US/09/904,621
;  CURRENT FILING DATE:  2001-07-13
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/265,627
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-09
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 5,455,338
;  PRIOR FILING DATE: EARLIER FILING DATE: 1993-11-05
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide ZC4792
US-09-904-621-3

  Query Match             4.2%;  Score 21;  DB 4;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 10
PCT-US93-11696-5
; Sequence 5, Application PC/TUS9311696
;  GENERAL INFORMATION:
;    APPLICANT:  Foster, Donald C
;    APPLICANT:  Sprecher, Cindy
;    APPLICANT:  Norris, Kjeld
;    TITLE OF INVENTION:  NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
;    TITLE OF INVENTION:  HOMOLOG AND KUNITZ-TYPE INHIBITOR
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  4225 Roosevelt Way, N.E.
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/11696
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/985,692
;       FILING DATE:  02-DEC-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Parker, Gary E
;       REGISTRATION NUMBER:  31-684
;       REFERENCE/DOCKET NUMBER:  92-21C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  206-547-8080 ext 322
;       TELEFAX:  206-548-2329
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  30 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       CLONE:  ZC4792
PCT-US93-11696-5

  Query Match             4.2%;  Score 21;  DB 5;  Length 30;
  Best Local Similarity   82.8%;  Pred. No. 4.8e+03;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              |||  |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29


RESULT 11
US-08-985-162-1712
; Sequence 1712, Application US/08985162
; Patent No. 6057156
;  GENERAL INFORMATION:
;    APPLICANT:  Akhtar, Saghir
;    APPLICANT:  Fell, Patricia
;    APPLICANT:  McSwiggen, James
;    TITLE OF INVENTION:  ENZYMATIC NUCLEIC ACID TREATMENT
;    TITLE OF INVENTION:  OF DISEASES OR CONDITIONS RELATED
;    TITLE OF INVENTION:  TO LEVELS OF EPIDERMAL GROWTH
;    TITLE OF INVENTION:  FACTOR RECEPTORS
;    NUMBER OF SEQUENCES:  1877
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,162
;      FILING DATE:  04 December 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/036,476
;      FILING DATE:  31 January 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  230/107
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213)  955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1712:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-985-162-1712

  Query Match             4.1%;  Score 20.2;  DB 3;  Length 50;
  Best Local Similarity   60.6%;  Pred. No. 9.3e+03;
  Matches   20;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        268 GCAACCAGAGGTTTGACATGGAAAAACATGTTG 300
              ||| ||||| |:::  ||  |||| ||| |::|
Db          1 GCACCCAGAAGUUUACCAGAGAAACACACGUUG 33


RESULT 12
US-09-401-063-1712
; Sequence 1712, Application US/09401063
; Patent No. 6623962
;  GENERAL INFORMATION:
;    APPLICANT:  Akhtar, Saghir
;    APPLICANT:  Fell, Patricia
;    APPLICANT:  McSwiggen, James
;    TITLE OF INVENTION:  ENZYMATIC NUCLEIC ACID TREATMENT
;    TITLE OF INVENTION:  OF DISEASES OR CONDITIONS RELATED
;    TITLE OF INVENTION:  TO LEVELS OF EPIDERMAL GROWTH
;    TITLE OF INVENTION:  FACTOR RECEPTORS
;    NUMBER OF SEQUENCES:  1877
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/401,063
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,162
;      FILING DATE:  04 December 1997
;      APPLICATION NUMBER:  60/036,476
;      FILING DATE:  31 January 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  230/107
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213)  955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1712:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-401-063-1712
```

Query Match 4.1%; Score 20.2; DB 4; Length 50;
Best Local Similarity 60.6%; Pred. No. 9.3e+03;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```
Qy        268 GCAACCAGAGGTTTGACATGGAAAAACATGTTG 300
              ||| ||||| |::: ||  |||| ||| |::|
Db          1 GCACCCAGAAGUUUACCAGAGAAACACACGUUG 33
```

RESULT 13
US-09-133-774-26
; Sequence 26, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636el Peptides Capable of Modulating Inflammatory Heart
; TITLE OF INVENTION: Disease
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Human
US-09-133-774-26

Query Match 4.0%; Score 19.8; DB 2; Length 42;
Best Local Similarity 77.4%; Pred. No. 1.1e+04;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy        115 CTGCAGCCTTTCTGCCACCTGTACTGGGGCT 145
              | ||| ||| | ||||| || | ||||||||
Db         12 CAGCACCCTGTTTGCCAACTATGCTGGGGCT 42
```

RESULT 14
US-09-303-862-26
; Sequence 26, Application US/09303862
; Patent No. 6034230
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 6034230el Peptides Capable of Modulating Inflammatory Heart
; TITLE OF INVENTION: Disease
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/303,862
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 09/133,774
; EARLIER FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Human
US-09-303-862-26

Query Match 4.0%; Score 19.8; DB 3; Length 42;
Best Local Similarity 77.4%; Pred. No. 1.1e+04;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy        115 CTGCAGCCTTTCTGCCACCTGTACTGGGGCT 145
              | ||| ||| | ||||| || | ||||||||
Db         12 CAGCACCCTGTTTGCCAACTATGCTGGGGCT 42
```

RESULT 15
US-08-985-162-1737
; Sequence 1737, Application US/08985162
; Patent No. 6057156
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir
; APPLICANT: Fell, Patricia
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; TITLE OF INVENTION: FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,162
; FILING DATE: 04 December 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/036,476
; FILING DATE: 31 January 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1737:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-985-162-1737

Query Match 4.0%; Score 19.8; DB 3; Length 50;
Best Local Similarity 51.1%; Pred. No. 1.2e+04;
Matches 24; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

```
Qy        224 TGAACAACAACAGCTACGAGTCAGACATCCTGAAGAATTACCTGGCA 270
              :||| || |||| |   ||  || || :  :|    |::|||:|| |
Db          4 UGAAGAAGAACACCAGAGAAACACACGUUGUGGUACAUUACCUGGUA 50
```

Search completed: May 26, 2004, 09:20:42
Job time : 97.1979 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:41:55 ; Search time 2173.84 Seconds
(without alignments)
1041.142 Million cell updates/sec

Title: US-10-048-046-1_COPY_1516_2013
Perfect score: 498
Sequence: 1 tgccctctgcaaggaagcca.....gctactggggccgtaactgc 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 1612472

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	4.2	30	9	US-09-904-621-3	Sequence 3, Appli
	2	20.8	4.2	48	12	US-10-251-085B-26	Sequence 26, Appl
	3	20.8	4.2	48	12	US-10-251-085B-64	Sequence 64, Appl
	4	20.4	4.1	47	16	US-10-415-247-6	Sequence 6, Appli
c	5	20	4.0	50	16	US-10-131-827-912	Sequence 912, App
c	6	19.6	3.9	48	11	US-09-767-764A-12	Sequence 12, Appl
c	7	19.6	3.9	48	15	US-10-285-010-20	Sequence 20, Appl
	8	19.6	3.9	50	16	US-10-131-827-7822	Sequence 7822, Ap
	9	19.4	3.9	39	15	US-10-173-461-24	Sequence 24, Appl
	10	19	3.8	36	15	US-10-008-960-26	Sequence 26, Appl
	11	18.8	3.8	48	12	US-10-251-085B-27	Sequence 27, Appl
	12	18.8	3.8	48	12	US-10-251-085B-38	Sequence 38, Appl
	13	18.8	3.8	48	12	US-10-251-085B-65	Sequence 65, Appl
	14	18.8	3.8	48	12	US-10-251-085B-76	Sequence 76, Appl
c	15	18.6	3.7	38	10	US-09-930-423-3110	Sequence 3110, Ap
c	16	18.6	3.7	38	10	US-09-745-237A-3110	Sequence 3110, Ap
	17	18.6	3.7	42	14	US-10-090-983-11	Sequence 11, Appl
	18	18.6	3.7	50	9	US-09-504-231A-2950	Sequence 2950, Ap
	19	18.6	3.7	50	9	US-09-504-231A-3045	Sequence 3045, Ap
	20	18.6	3.7	50	9	US-09-274-553D-2950	Sequence 2950, Ap
	21	18.6	3.7	50	9	US-09-274-553D-3045	Sequence 3045, Ap
	22	18.6	3.7	50	16	US-10-131-827-2864	Sequence 2864, Ap
	23	18.4	3.7	36	15	US-10-205-942-25	Sequence 25, Appl
c	24	18.4	3.7	36	15	US-10-205-942-26	Sequence 26, Appl
	25	18.4	3.7	37	17	US-10-138-674-18176	Sequence 18176, A
c	26	18.4	3.7	41	12	US-10-035-833A-4680	Sequence 4680, Ap
	27	18.4	3.7	45	9	US-09-765-873A-29	Sequence 29, Appl
c	28	18.4	3.7	45	9	US-09-765-873A-30	Sequence 30, Appl
	29	18.4	3.7	45	15	US-10-188-523B-29	Sequence 29, Appl
c	30	18.4	3.7	45	15	US-10-188-523B-30	Sequence 30, Appl
	31	18.4	3.7	47	16	US-10-349-143-1087	Sequence 1087, Ap
	32	18.4	3.7	50	16	US-10-131-827-390	Sequence 390, App
	33	18.4	3.7	50	16	US-10-131-827-7211	Sequence 7211, Ap
c	34	18.2	3.7	46	9	US-09-766-898-14	Sequence 14, Appl
c	35	18.2	3.7	46	9	US-09-766-898-25	Sequence 25, Appl
c	36	18.2	3.7	46	9	US-09-766-916-14	Sequence 14, Appl
c	37	18.2	3.7	46	9	US-09-766-916-25	Sequence 25, Appl
c	38	18.2	3.7	46	15	US-10-306-247-14	Sequence 14, Appl
c	39	18.2	3.7	46	15	US-10-306-247-25	Sequence 25, Appl
c	40	18.2	3.7	46	15	US-10-306-246-14	Sequence 14, Appl
c	41	18.2	3.7	46	15	US-10-306-246-25	Sequence 25, Appl
c	42	18.2	3.7	46	15	US-10-306-249-14	Sequence 14, Appl
c	43	18.2	3.7	46	15	US-10-306-249-25	Sequence 25, Appl
c	44	18.2	3.7	46	15	US-10-306-013-14	Sequence 14, Appl
c	45	18.2	3.7	46	15	US-10-306-013-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-904-621-3
; Sequence 3, Application US/09904621
; Patent No. US20020098560A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kisiel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS
; TITLE OF INVENTION: AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; FILE REFERENCE: 93-14D3
; CURRENT APPLICATION NUMBER: US/09/904,621
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/265,627
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 5,455,338
; PRIOR FILING DATE: EARLIER FILING DATE: 1993-11-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide ZC4792
US-09-904-621-3

Query Match 4.2%; Score 21; DB 9; Length 30;
Best Local Similarity 82.8%; Pred. No. 1.1e+04;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        377 GTTACTGCTGTGGCCTGCGCAGCTTCCGT 405
              ||| |||||| |||| |||||| |||||
Db          1 GTTGTTGCTGTTGCCTCCGCAGCCTCCGT 29
```

```
RESULT 2
US-10-251-085B-26
; Sequence 26, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  APPLICANT:  Lin, Ying-Chi
;  APPLICANT:  Maruyama, Toshiaki
;  TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
;  FILE REFERENCE: 1087-21
;  CURRENT APPLICATION NUMBER: US/10/251,085B
;  CURRENT FILING DATE:  2002-09-19
;  PRIOR APPLICATION NUMBER: US 60/323,455
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-10-251-085B-26

  Query Match             4.2%;  Score 20.8;  DB 12;  Length 48;
  Best Local Similarity   70.0%;  Pred. No. 1.5e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || ||||||||| || || |||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGCAGGTCCAGCTTGT 44


RESULT 3
US-10-251-085B-64
; Sequence 64, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  APPLICANT:  Lin, Ying-Chi
;  APPLICANT:  Maruyama, Toshiaki
;  TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
;  FILE REFERENCE: 1087-21
;  CURRENT APPLICATION NUMBER: US/10/251,085B
;  CURRENT FILING DATE:  2002-09-19
;  PRIOR APPLICATION NUMBER: US 60/323,455
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 64
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-10-251-085B-64

  Query Match             4.2%;  Score 20.8;  DB 12;  Length 48;
  Best Local Similarity   70.0%;  Pred. No. 1.5e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || ||||||||| || || |||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGCAGGTCCAGCTTGT 44


RESULT 4
US-10-415-247-6
; Sequence 6, Application US/10415247
; Publication No. US20040013655A1
; GENERAL INFORMATION:
;  APPLICANT: Shiozawa, Shunichi
;  TITLE OF INVENTION: Genome responsible for chronic rheumatoid arthritis,
;  TITLE OF INVENTION:  diagnostic method, pathogenicity judging method and
;  TITLE OF INVENTION:  detection-use diagnostic kit of chronic rheumatoid
;  TITLE OF INVENTION:  arthritis, and therapeutic method and medicine of
;  TITLE OF INVENTION:  chronic rheumatoid arthritis
;  FILE REFERENCE: TLOP1-2
;  CURRENT APPLICATION NUMBER: US/10/415,247
;  CURRENT FILING DATE:  2003-04-24
;  PRIOR APPLICATION NUMBER: JP 2000-324296
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: JP 2001-90546
;  PRIOR FILING DATE: 2001-3-27
;  PRIOR APPLICATION NUMBER: JP 2001-99990
;  PRIOR FILING DATE: 2001-3-30
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthesized oligonucleotide
US-10-415-247-6

  Query Match             4.1%;  Score 20.4;  DB 16;  Length 47;
  Best Local Similarity   65.2%;  Pred. No. 2e+04;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         27 GTGCACCTGCTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGC 72
              | | |||  | |||||| ||||  ||| | | ||   | |||| ||
Db          2 GGGTACCATCCGCTTCCTGCCCCAGCCAGGCTGGTTTGTGGAGTGC 47


RESULT 5
US-10-131-827-912/c
; Sequence 912, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 912
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-131-827-912

  Query Match             4.0%;  Score 20;  DB 16;  Length 50;
  Best Local Similarity   72.2%;  Pred. No. 2.7e+04;
  Matches   26;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        198 GGGTGACAAGTGTCTGGACGGCGTGCTGAACAACAA 233
              |||  | |||||| |||| | | |||  ||||| ||
Db         41 GGGGCAGAAGTGTGTGGATGACCTGCCAAACAATAA 6


RESULT 6
US-09-767-764A-12/c
```

; Sequence 12, Application US/09767764A
; Publication No. US20030207265A1
; GENERAL INFORMATION:
; APPLICANT: Church, George M.
; TITLE OF INVENTION: METHOD OF MAKING PROTEIN ARRAYS
; FILE REFERENCE: 10498-00009
; CURRENT APPLICATION NUMBER: US/09/767,764A
; CURRENT FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/522,732
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Amplification primer
US-09-767-764A-12

Query Match 3.9%; Score 19.6; DB 11; Length 48;
Best Local Similarity 73.5%; Pred. No. 3.5e+04;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 36 CTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAG 69
|| |||| ||| ||||| ||| |||||||
Db 43 CTCCTTCAAGCTTCCGCCCGAGAGGAAAGCGGAG 10

RESULT 7
US-10-285-010-20/c
; Sequence 20, Application US/10285010
; Publication No. US20030124594A1
; GENERAL INFORMATION:
; APPLICANT: Church, George
; APPLICANT: Mitra, Robi
; TITLE OF INVENTION: Replica Amplification Of Nucleic Acid Arrays
; FILE REFERENCE: 10498-00040
; CURRENT APPLICATION NUMBER: US/10/285,010
; CURRENT FILING DATE: 2002-10-31
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amplification primer
US-10-285-010-20

Query Match 3.9%; Score 19.6; DB 15; Length 48;
Best Local Similarity 73.5%; Pred. No. 3.5e+04;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 36 CTGCTTCCAGCCCATGCCCGACCGGAGAGCGGAG 69
|| |||| ||| ||||| ||| |||||||
Db 43 CTCCTTCAAGCTTCCGCCCGAGAGGAAAGCGGAG 10

RESULT 8
US-10-131-827-7822
; Sequence 7822, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7822
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-7822

Query Match 3.9%; Score 19.6; DB 16; Length 50;
Best Local Similarity 66.7%; Pred. No. 3.6e+04;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 197 TGGGTGACAAGTGTCTGGACGGCGTGCTGAACAACAACAGCT 238
||| | ||| |||||| || | |||| ||||| || |
Db 1 TGGCGAGCCAGTCTCTGGATGGGATTCTGATCAACAGAAGTT 42

RESULT 9
US-10-173-461-24
; Sequence 24, Application US/10173461
; Publication No. US20030138795A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN GROWTH FACTOR WITH HOMOLOGY
; TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR, BGS-8, EXPRESSED HIGHLY IN IMMUNE TISSUE
; FILE REFERENCE: D0166 NP
; CURRENT APPLICATION NUMBER: US/10/173,461
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/298,340
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-461-24

Query Match 3.9%; Score 19.4; DB 15; Length 39;
Best Local Similarity 70.3%; Pred. No. 3.9e+04;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 158 GCTGCTACGGCTGCCTGGCCCCGTTTTGTGAGCTCAA 194
|| || |||| || ||||||| | |||||| | |
Db 1 GCAGCAGCGGCCGCATGGCCCCCTGGAGTGAGCACCA 37

RESULT 10
US-10-008-960-26
; Sequence 26, Application US/10008960
; Publication No. US20030133971A1
; GENERAL INFORMATION:
; APPLICANT: BAYLOR COLLEGE OF MEDICINE
; APPLICANT: SMITH, JAMES R.
; APPLICANT: DRUTZ, DAVID J.
; APPLICANT: WILSON, DEBORAH R.
; APPLICANT: ZUMSTEIN, LOUIS A.
; TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
; TITLE OF INVENTION: DNA SYNTHESIS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROGERS & WELLS
; STREET: 200 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10166
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/10/008,960
;       FILING DATE:  07-DEC-2001
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/327,874
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/808,523
;       FILING DATE:  16-DEC-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/970,462
;       FILING DATE:  02-NOV-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/113,372
;       FILING DATE:  30-AUG-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/153,564
;       FILING DATE:  17-NOV-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/203,535
;       FILING DATE:  25-FEB-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/229,420
;       FILING DATE:  15-APR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/274,535
;       FILING DATE:  13-JUL-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT US94/09700
;       FILING DATE:  26-AUG-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/327,874
;       FILING DATE:  24-OCT-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  NORTON, GERARD P.
;       REGISTRATION NUMBER:  36,621
;       REFERENCE/DOCKET NUMBER:  3634-8-CIP10
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 878-3148
;       TELEFAX:  (212) 878-8375
;   INFORMATION FOR SEQ ID NO:  26:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  36 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     IMMEDIATE SOURCE:
;       CLONE:  Primer
US-10-008-960-26

  Query Match             3.8%;  Score 19;  DB 15;  Length 36;
  Best Local Similarity   71.4%;  Pred. No. 5e+04;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        458 CCGTAACATCCCGTCCTGACTGCTACTGGGGCCGT 492
              |||  ||||||||  | ||||  |  |||||  |||
Db          2 CCGAGACATCCCGGGCCGACTTCGCCTGGGAGCGT 36


RESULT 11
US-10-251-085B-27
; Sequence 27, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  APPLICANT:  Lin, Ying-Chi
;  APPLICANT:  Maruyama, Toshiaki
;  TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
;  FILE REFERENCE: 1087-21
;  CURRENT APPLICATION NUMBER: US/10/251,085B
;  CURRENT FILING DATE:  2002-09-19
;  PRIOR APPLICATION NUMBER: US 60/323,455
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-10-251-085B-27

  Query Match             3.8%;  Score 18.8;  DB 12;  Length 48;
  Best Local Similarity   65.0%;  Pred. No. 6.4e+04;
  Matches   26;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || ||||:||| || ||  ||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGSAGGTCCAGCTGGT 44


RESULT 12
US-10-251-085B-38
; Sequence 38, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  APPLICANT:  Lin, Ying-Chi
;  APPLICANT:  Maruyama, Toshiaki
;  TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
;  FILE REFERENCE: 1087-21
;  CURRENT APPLICATION NUMBER: US/10/251,085B
;  CURRENT FILING DATE:  2002-09-19
;  PRIOR APPLICATION NUMBER: US 60/323,455
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-10-251-085B-38

  Query Match             3.8%;  Score 18.8;  DB 12;  Length 48;
  Best Local Similarity   65.0%;  Pred. No. 6.4e+04;
  Matches   26;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || |||||||| :| ||  ||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGCAGGTSCAGCTGGT 44


RESULT 13
US-10-251-085B-65
; Sequence 65, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
;  APPLICANT: Bowdish, Katherine S.
;  APPLICANT:  Frederickson, Shana
;  APPLICANT:  Renshaw, Mark
;  APPLICANT:  Lin, Ying-Chi
;  APPLICANT:  Maruyama, Toshiaki
```

; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21
; CURRENT APPLICATION NUMBER: US/10/251,085B
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 65
; LENGTH: 48
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
US-10-251-085B-65

Query Match 3.8%; Score 18.8; DB 12; Length 48;
Best Local Similarity 65.0%; Pred. No. 6.4e+04;
Matches 26; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

```
Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || ||||:||| || ||  ||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGSAGGTCCAGCTGGT 44
```

RESULT 14
US-10-251-085B-76
; Sequence 76, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Maruyama, Toshiaki
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21
; CURRENT APPLICATION NUMBER: US/10/251,085B
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 76
; LENGTH: 48
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
US-10-251-085B-76

Query Match 3.8%; Score 18.8; DB 12; Length 48;
Best Local Similarity 65.0%; Pred. No. 6.4e+04;
Matches 26; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

```
Qy         50 TGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTGT 89
              || |||   | ||||  | || ||||||||:|| ||  ||
Db          5 TGGCCGTTGGAAGAGGAGTGCTCGAGCAGGTSCAGCTGGT 44
```

RESULT 15
US-09-930-423-3110/c
; Sequence 3110, Application US/09930423
; Publication No. US20030092003A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: MBHB00,918-A 400/027
; CURRENT APPLICATION NUMBER: US/09/930,423
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 4553
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3110
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-930-423-3110

Query Match 3.7%; Score 18.6; DB 10; Length 38;
Best Local Similarity 72.7%; Pred. No. 6.9e+04;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy         64 GCGGAGCGCGAGCAGGACCCGCGTGTCGCCCCT 96
              ||| |||| || |  ||| ||| | ||| ||||
Db         38 GCGCAGCGAGACCTTGACTCGCCTTTCGGCCCT 6
```

Search completed: May 26, 2004, 10:22:40
Job time : 2173.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:12:53 ; Search time 3457.26 Seconds
(without alignments)
4301.492 Million cell updates/sec

Title: US-10-048-046-1_COPY_1516_2013
Perfect score: 498
Sequence: 1 tgcctctgcaaggaagcca.....gctactggggccgtaactgc 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20.8	4.2	50	9	AU105995	AU105995 AU105995
2	20.8	4.2	50	9	AU105997	AU105997 AU105997
3	20.6	4.1	50	9	AU105996	AU105996 AU105996
4	20.2	4.1	47	28	CC178014	CC178014 XC898 Bay
c 5	19.6	3.9	43	9	AA974059	AA974059 oq05h07.s
c 6	19.6	3.9	50	9	AU103246	AU103246 AU103246
c 7	19.4	3.9	49	28	AZ606700	AZ606700 1M0428E15
c 8	19.2	3.9	40	9	AA990168	AA990168 ua60e08.r
c 9	19	3.8	49	9	AI271755	AI271755 qj77c07.x
10	18.8	3.8	50	9	AU105998	AU105998 AU105998
c 11	18.6	3.7	40	10	BF343276	BF343276 602015935
c 12	18.6	3.7	49	9	AA921726	AA921726 oh08d02.s
c 13	18.6	3.7	50	9	AU103244	AU103244 AU103244
14	18.6	3.7	50	9	AU104470	AU104470 AU104470
c 15	18.6	3.7	50	9	AU105005	AU105005 AU105005
16	18.6	3.7	50	12	BI665523	BI665523 603289585
17	18.4	3.7	46	9	AI949114	AI949114 wq19g02.x
c 18	18.4	3.7	50	9	AU105237	AU105237 AU105237
c 19	18.4	3.7	50	9	AU107915	AU107915 AU107915
c 20	18.2	3.7	48	29	CG779308	CG779308 1123033B1
c 21	18.2	3.7	49	9	AA933611	AA933611 oh87e12.s
22	18	3.6	37	9	AI048978	AI048978 uc77d09.y
23	18	3.6	43	12	BI103457	BI103457 602888987
c 24	18	3.6	46	28	AZ512864	AZ512864 1M0358A13
c 25	18	3.6	50	9	AU103249	AU103249 AU103249
26	18	3.6	50	9	AU105753	AU105753 AU105753
27	18	3.6	50	9	AU105754	AU105754 AU105754
28	18	3.6	50	9	AU105755	AU105755 AU105755
c 29	18	3.6	50	9	AU107902	AU107902 AU107902
c 30	18	3.6	50	9	AU107933	AU107933 AU107933
c 31	17.8	3.6	38	28	AZ868035	AZ868035 2M0179C07
c 32	17.8	3.6	46	9	AA879784	AA879784 vv98f04.r
33	17.8	3.6	46	28	CC183613	CC183613 RST267 Ba
34	17.8	3.6	49	28	BH627713	BH627713 1007074H1
35	17.8	3.6	50	9	AU102353	AU102353 AU102353
36	17.8	3.6	50	9	AU102354	AU102354 AU102354
c 37	17.8	3.6	50	9	AU105314	AU105314 AU105314
38	17.8	3.6	50	9	AU106718	AU106718 AU106718
c 39	17.6	3.5	46	9	AI688909	AI688909 tx22e06.x
40	17.6	3.5	46	9	AA410197	AA410197 zv32d06.r
c 41	17.6	3.5	49	9	AI001027	AI001027 os49h01.s
42	17.6	3.5	49	12	BM862480	BM862480 mgcm003xC
c 43	17.6	3.5	50	9	AU103245	AU103245 AU103245
44	17.6	3.5	50	9	AU104468	AU104468 AU104468
45	17.6	3.5	50	9	AU105698	AU105698 AU105698

ALIGNMENTS

RESULT 1
AU105995
LOCUS AU105995 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU105995 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone ADSE00435, mRNA sequence.
ACCESSION AU105995
VERSION AU105995.1 GI:13555516
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and

```
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ADSE00435"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             4.2%;  Score 20.8;  DB 9;  Length 50;
  Best Local Similarity   64.6%;  Pred. No. 6.5e+05;
  Matches   31;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         38 GCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGC 85
              | |  | ||| | || || |||||| | ||||| || || |   || |
Db          1 GTTGGCTGCCGAGGCTCGGCCGGAGCGTGGAGCCCGCGCCGCTGCCCC 48


RESULT 2
AU105997
LOCUS       AU105997                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105997 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
          KAIA0410, mRNA sequence.
ACCESSION   AU105997
VERSION     AU105997.1  GI:13555518
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
          Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
          Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
          mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAIA0410"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             4.2%;  Score 20.8;  DB 9;  Length 50;
  Best Local Similarity   64.6%;  Pred. No. 6.5e+05;
  Matches   31;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         38 GCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGC 85
              | |  | ||| | || || |||||| | ||||| || || |   || |
Db          1 GTTGGCTGCCGAGGCTCGGCCGGAGCGTGGAGCCCGCGCCGCTGCCCC 48


RESULT 3
AU105996
LOCUS       AU105996                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105996 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
          ADSU02176, mRNA sequence.
ACCESSION   AU105996
VERSION     AU105996.1  GI:13555517
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
          Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
          Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
          mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
          Sugano,S. Construction and characterization of a full
          length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
          149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ADSU02176"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             4.1%;  Score 20.6;  DB 9;  Length 50;
  Best Local Similarity   67.4%;  Pred. No. 7.3e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         43 CAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGC 85
              | ||| | || || |||||| | ||||| || || |   || |
Db          6 CTGCCGAGGCTCGGCCGGAGCGTGGAGCCCGCGCCGCTGCCCC 48


RESULT 4
CC178014
LOCUS       CC178014                 47 bp    mRNA    linear   GSS 02-MAY-2003
DEFINITION  XC898 BayGenomics Gene Trap Library pGT1Lxf Mus musculus cDNA, mRNA
          sequence.
ACCESSION   CC178014
VERSION     CC178014.1  GI:30316565
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   BayGenomics.
  TITLE     http://baygenomics.ucsf.edu/
  JOURNAL   Unpublished (2001)
COMMENT     Contact: BayGenomics
          Bay Area Functional Genomics Consortium (BayGenomics)
          Email: info@baygenomics.ucsf.edu
          Sequence tag generated by 5' RACE of total RNA from gene trap ES
          cell line. ES cell lines harboring insertion mutation of target
          gene are available upon request from BayGenomics. Annotation
          information available from
          http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=
          CELL_LINE&KEY=XC898
          Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Mus musculus"
```

```
                     /mol_type="mRNA"
                     /strain="129 OLA"
                     /db_xref="taxon:10090"
                     /sex="Male"
                     /cell_type="Embryonic stem cell"
                     /clone_lib="BayGenomics Gene Trap Library pGT1Lxf"
                     /note="Vector: pGT1Lxf"
ORIGIN

  Query Match             4.1%;  Score 20.2;  DB 28;  Length 47;
  Best Local Similarity   68.3%;  Pred. No. 8.9e+05;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        113 GCCTGCAGCCTTTCTGCCACCTGTACTGGGGCTGCACCCGG 153
              || ||| || |||| ||| | |   ||| |||| |||| |
Db          7 GCTTGCTGCTTTTCCGCCCTCCGAGATGGCGCTGGACCCCG 47


RESULT 5
AA974059/c
LOCUS       AA974059                 43 bp    mRNA    linear   EST 07-JUL-1998
DEFINITION  oq05h07.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1585501 3'
            similar to SW:ASH1_HUMAN P50553 ACHAETE-SCUTE HOMOLOG 1. ;contains
            element TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AA974059
VERSION     AA974059.1  GI:3149239
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 802   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1585501"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu5"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     neuroendocrine lung carcinoid, and was then primed with a
                     Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                     to Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector.  Library is normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             3.9%;  Score 19.6;  DB 9;  Length 43;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+06;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        146 GCACCCGGACCGGCTGCTACGGCTGCCTGGCCCCGTTTTGTG 187
              || ||| || |||||||  |||| ||| | | | | | | |
Db         42 GCCCCCTGAGCGGCTGCGTCGGCCGCCGGTCTCAGCTCGGCG 1


RESULT 6
AU103246/c
LOCUS       AU103246                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU103246 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            KAT01774, mRNA sequence.
ACCESSION   AU103246
VERSION     AU103246.1  GI:13552767
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAT01774"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             3.9%;  Score 19.6;  DB 9;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 1.3e+06;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         40 TTCCAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGAC 81
              | |||| ||  ||||   | |||  |||  || |||| ||||
Db         45 TACCAGACCGCGCCCTTGCTGAGGACGGCCCGGGAGCCGGAC 4


RESULT 7
AZ606700/c
LOCUS       AZ606700                 49 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0428E15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0428E15 R, genomic survey sequence.
ACCESSION   AZ606700
VERSION     AZ606700.1  GI:11728890
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
```

```
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0428  row: E  column: 15
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0428E15"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             3.9%;  Score 19.4;  DB 28;  Length 49;
  Best Local Similarity   64.4%;  Pred. No. 1.4e+06;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          90 CGCCCCTCAGCAGTGTGCGGTCTGCCTGCAGCCTTTCTGCCACCT 134
               | |||||    |   ||||| || |  |||||||| || | || |
Db          47 CTCCCCTTCTTACCCTGCGGGCTTCAGGCAGCCTTCCTCCAACAT 3


RESULT 8
AA990168/c
LOCUS       AA990168                 40 bp    mRNA    linear   EST 02-JUN-1998
DEFINITION  ua60e08.r1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:1361894 5' similar to TR:O14731 O14731 MEMBRANE-ASSOCIATED
            KINASE. [1] ;, mRNA sequence.
ACCESSION   AA990168
VERSION     AA990168.1  GI:3175532
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:699934
            Trace considered overall poor quality
            Putative full length read
            vector to vector length is 532
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1361894"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_thymus_2NbMT"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             3.9%;  Score 19.2;  DB 9;  Length 40;
  Best Local Similarity   67.5%;  Pred. No. 1.4e+06;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          43 CAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGACC 82
               ||||| |||| || |    || | || | || |||| ||
Db          40 CAGCCCGTGCCAGAGCCAGCAGAGAAGTGTGACCAGGGCC 1


RESULT 9
AI271755/c
LOCUS       AI271755                 49 bp    mRNA    linear   EST 17-MAR-1999
DEFINITION  qj77c07.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1865484 3'
            similar to TR:O00411 O00411 MITOCHONDRIAL RNA POLYMERASE. ;contains
            MER22.b3 TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AI271755
VERSION     AI271755.1  GI:3890922
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
```

cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 1806 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

```
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1865484"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid3"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo. "
ORIGIN

  Query Match             3.8%;  Score 19;  DB 9;  Length 49;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+06;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         62 GAGCGGAGCGCGAGCAGGACCCGCGTGTCGCCCCT 96
              | ||| | | | |||||| ||||||  ||| | ||
Db         39 GGGCGCAACTCCAGCAGGGCCCGCGCCTCGTCGCT 5


RESULT 10
AU105998
LOCUS       AU105998                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105998 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            W0521, mRNA sequence.
ACCESSION   AU105998
VERSION     AU105998.1  GI:13555519
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="W0521"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             3.8%;  Score 18.8;  DB 9;  Length 50;
  Best Local Similarity   63.0%;  Pred. No. 2e+06;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         38 GCTTCCAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCC 83
              | |  | ||| | || || |||||| | ||||| ||  |    |||
Db          1 GTTGGCTGCCGAGGCTCGGCCGGAGCGTGGAGCCCGGCCGCTGCCC 46


RESULT 11
BF343276/c
LOCUS       BF343276                 40 bp    mRNA    linear   EST 22-NOV-2000
DEFINITION  602015935F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4151432
            5', mRNA sequence.
ACCESSION   BF343276
VERSION     BF343276.1  GI:11290436
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9416 row: f column: 09
            High quality sequence stop: 38.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4151432"
                     /tissue_type="glioblastoma with EGFR amplification"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Brn64"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.57 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             3.7%;  Score 18.6;  DB 10;  Length 40;
  Best Local Similarity   72.7%;  Pred. No. 2e+06;
  Matches   24;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy        140 GGGGCTGCACCCGGACCGGCTGCTACGGCTGCC 172
              || ||||||  || ||| ||||  || || |||
Db         33 GGAGCTGCAGACGCACCCGCTGGGACCGCAGCC 1


RESULT 12
AA921726/c
LOCUS       AA921726                 49 bp    mRNA    linear   EST 13-MAY-1998
DEFINITION  oh08d02.s1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1457187 3'
            similar to TR:O15551 O15551 RAT VENTRAL PROSTATE.1 HOMOLOG. ;, mRNA
            sequence.
ACCESSION   AA921726
VERSION     AA921726.1  GI:3068914
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 739   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1457187"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Co8"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             3.7%;  Score 18.6;  DB 9;  Length 49;
  Best Local Similarity   84.0%;  Pred. No. 2.2e+06;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         21 CGCCCTGTGCACCTGCTGCTTCCAG 45
              ||||| |||||||| |||| | |||
Db         35 CGCCCAGTGCACCTACTGCGTGCAG 11


RESULT 13
AU103244/c
LOCUS       AU103244                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU103244 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            COLF6658, mRNA sequence.
ACCESSION   AU103244
VERSION     AU103244.1  GI:13552765
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COLF6658"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             3.7%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   61.2%;  Pred. No. 2.3e+06;
  Matches   30;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         40 TTCCAGCCCATGCCCGACCGGAGAGCGGAGCGCGAGCAGGACCCGCGTG 88
              | |||| ||  ||||   | |||  | |  || |||| ||||    |||
Db         49 TACCAGACCGCGCCCTTGCTGAGGACAGCCCGGGAGCCGGACAGCGGTG 1


RESULT 14
AU104470
LOCUS       AU104470                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104470 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            KAT02069, mRNA sequence.
ACCESSION   AU104470
VERSION     AU104470.1  GI:13553991
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAT02069"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             3.7%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   65.9%;  Pred. No. 2.3e+06;
  Matches   27;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy        303 CGAGAGCCTCGTGGCTCTCCAGCGGGGAGTGTTTCTGCTGT 343
              ||||  |  | |||| | | ||| | | | | | |||||||
```

```
Db          1 CGAGTTCGACATGGCGCGCGAGCTGCGGGCGCTGCTGCTGT 41


RESULT 15
AU105005/c
LOCUS       AU105005                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105005 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            ADSE01748, mRNA sequence.
ACCESSION   AU105005
VERSION     AU105005.1  GI:13554526
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ADSE01748"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             3.7%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   65.9%;  Pred. No. 2.3e+06;
  Matches   27;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         58 CGGAGAGCGGAGCGCGAGCAGGACCCGCGTGTCGCCCCTCA 98
              ||| |  |  ||  |   ||| ||||||| |||||| || |
Db         41 CGGCGTTCTAAGGCCTCCCAGCACCCGCGCGTCGCCGCTTA 1


Search completed: May 26, 2004, 09:17:50
Job time : 3459.26 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:10:28 ; Search time 1630.25 Seconds
(without alignments)
5822.506 Million cell updates/sec

Title: US-10-048-046-1_COPY_181_399
Perfect score: 219
Sequence: 1 gtcctcctgaggaagcggga.....tacagactggggatgtcatc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		19.8	9.0	50	6	AX165857	AX165857 Sequence
2		19.4	8.9	40	6	A83617	A83617 Sequence 46
3	c	19	8.7	47	6	AX194693	AX194693 Sequence
4	c	18.8	8.6	38	6	AR013934	AR013934 Sequence
5	c	18.4	8.4	38	6	AR020204	AR020204 Sequence
6	c	18.4	8.4	38	6	AR032156	AR032156 Sequence
7	c	18.4	8.4	38	6	AR043821	AR043821 Sequence
8	c	18.4	8.4	38	6	AR072133	AR072133 Sequence
9	c	18.4	8.4	38	6	AR095987	AR095987 Sequence
10	c	18.4	8.4	38	6	AR136647	AR136647 Sequence
11	c	18.4	8.4	38	6	AR157182	AR157182 Sequence
12	c	18.4	8.4	38	6	AR164131	AR164131 Sequence
13	c	18.4	8.4	38	6	I32817	I32817 Sequence 13
14	c	18.4	8.4	38	6	I59631	I59631 Sequence 13
15	c	18.4	8.4	38	6	AX127254	AX127254 Sequence
16	c	18.4	8.4	38	6	AX800785	AX800785 Sequence
17		18.4	8.4	46	6	I18512	I18512 Sequence 12
18		18.4	8.4	47	6	AR289961	AR289961 Sequence
19	c	18.2	8.3	43	6	BD174678	BD174678 Ribozyme
20		18	8.2	29	6	BD140120	BD140120 Secreted
21		18	8.2	42	6	AX328821	AX328821 Sequence
22		18	8.2	42	6	BD132386	BD132386 DNA diagn
23		18	8.2	50	6	E44269	E44269 Oligo-DNA s
24		18	8.2	50	6	AX697058	AX697058 Sequence
25	c	17.8	8.1	49	6	AR032442	AR032442 Sequence
26	c	17.8	8.1	49	6	I29182	I29182 Sequence 54
27	c	17.8	8.1	49	6	I90856	I90856 Sequence 54
28	c	17.8	8.1	49	6	AR209106	AR209106 Sequence
29		17.6	8.0	40	6	AX060727	AX060727 Sequence
30		17.6	8.0	40	6	AX060906	AX060906 Sequence
31	c	17.6	8.0	41	6	AX518097	AX518097 Sequence
32	c	17.4	7.9	41	6	AX513875	AX513875 Sequence
33	c	17.4	7.9	41	6	AX519151	AX519151 Sequence
34		17.4	7.9	47	6	AR284501	AR284501 Sequence
35		17.4	7.9	50	6	AR032934	AR032934 Sequence
36		17.4	7.9	50	6	I29674	I29674 Sequence 54
37		17.4	7.9	50	6	I91348	I91348 Sequence 54
38		17.4	7.9	50	6	AR209598	AR209598 Sequence
39		17.2	7.9	38	6	AR210306	AR210306 Sequence
40		17.2	7.9	38	6	AX402734	AX402734 Sequence
41	c	17.2	7.9	41	6	AR061380	AR061380 Sequence
42	c	17.2	7.9	41	6	AR108279	AR108279 Sequence
43	c	17.2	7.9	41	6	I16236	I16236 Sequence 62
44	c	17.2	7.9	41	6	I66722	I66722 Sequence 62
45	c	17.2	7.9	41	6	I84816	I84816 Sequence 62

ALIGNMENTS

RESULT 1
AX165857
LOCUS AX165857 50 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1052 from Patent WO0138586.
ACCESSION AX165857
VERSION AX165857.1 GI:14546686
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof

```
  JOURNAL   Patent: WO 0138586-A 1052 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg44004690"
     variation       26
                     /note="single nucleotide polymorphism"
ORIGIN

  Query Match             9.0%;  Score 19.8;  DB 6;  Length 50;
  Best Local Similarity   69.2%;  Pred. No. 2.3e+05;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          9 GAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGTTG 47
              ||||| | ||  |||||  |    || ||||| |||| ||
Db          1 GAGGAGGAGGAGGTGGAGGAGGAGGGGGAGAAGAGGATG 39


RESULT 2
A83617
LOCUS       A83617                    40 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 46 from Patent WO9849324.
ACCESSION   A83617
VERSION     A83617.1  GI:6732869
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Matthijs,G.
  TITLE     CARBOHYDRATE-DEFICIENT GLYCOPROTEIN SYNDROME TYPE I
  JOURNAL   Patent: WO 9849324-A 46 05-NOV-1998;
            MATTHIJS GERT (BE); GENZYME LTD (GB)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             8.9%;  Score 19.4;  DB 6;  Length 40;
  Best Local Similarity   79.3%;  Pred. No. 3.1e+05;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         97 ATTGTAGTGGATGAAAAATCAGGTCAGGT 125
              |||||||||| |||||| |    | |||||
Db          4 ATTGTAGTGGTTGAAAAGTGTAGACAGGT 32


RESULT 3
AX194693/c
LOCUS       AX194693                  47 bp    DNA     linear   PAT 28-AUG-2001
DEFINITION  Sequence 163 from Patent WO0151659.
ACCESSION   AX194693
VERSION     AX194693.1  GI:15385340
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chu,T., Blumenfeld,M. and Cohen,D.
  TITLE     Biallelic markers derived from genomic regions carrying genes
            involved in central nervous system disorders
  JOURNAL   Patent: WO 0151659-A 163 19-JUL-2001;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     variation       24
                     /note="99-27562-366 : polymorphic base G or T"
ORIGIN

  Query Match             8.7%;  Score 19;  DB 6;  Length 47;
  Best Local Similarity   67.6%;  Pred. No. 4.2e+05;
  Matches   25;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        170 TGAAGGTTGTTAAGAAGCAGACATGCCCTTTACAGAC 206
              || ||| |||| | :||||  | || ||| |||  ||
Db         38 TGTAGGATGTTTAAMAGCATTCCTGGCCTCTACCCAC 2


RESULT 4
AR013934/c
LOCUS       AR013934                  38 bp    DNA     linear   PAT 05-DEC-1998
DEFINITION  Sequence 23 from patent US 5773223.
ACCESSION   AR013934
VERSION     AR013934.1  GI:3971388
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Shyamala,V. and Olson,P.Tekamp.
  TITLE     Endothelin B.sub.1, (ETB.sub.1) receptor polypeptide and its
            encoding nucleic acid methods, and uses thereof
  JOURNAL   Patent: US 5773223-A 23 30-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.6%;  Score 18.8;  DB 6;  Length 38;
  Best Local Similarity   68.4%;  Pred. No. 4.9e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         65 GCAATAAACTGGTCTCTGGAGATCACTGTAGAATTGTA 102
              |||| |    | | |||| |||| |   ||||||||||
Db         38 GCAACAGCTCGATATCTGAAGATAAAAATAGAATTGTA 1


RESULT 5
AR020204/c
LOCUS       AR020204                  38 bp    DNA     linear   PAT 05-DEC-1998
DEFINITION  Sequence 13 from patent US 5789156.
ACCESSION   AR020204
VERSION     AR020204.1  GI:3974819
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Tetracycline-regulated transcriptional inhibitors
  JOURNAL   Patent: US 5789156-A 13 04-AUG-1998;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3
```

```
RESULT 6
AR032156/c
LOCUS       AR032156                38 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 13 from patent US 5866755.
ACCESSION   AR032156
VERSION     AR032156.1  GI:5946445
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Animals transgenic for a tetracycline-regulated transcriptional
            inhibitor
  JOURNAL   Patent: US 5866755-A 13 02-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 7
AR043821/c
LOCUS       AR043821                38 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 13 from patent US 5814618.
ACCESSION   AR043821
VERSION     AR043821.1  GI:5964829
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Methods for regulating gene expression
  JOURNAL   Patent: US 5814618-A 13 29-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 8
AR072133/c
LOCUS       AR072133                38 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 13 from patent US 5912411.
ACCESSION   AR072133
VERSION     AR072133.1  GI:7223021
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Mice transgenic for a tetracycline-inducible transcriptional
            activator
  JOURNAL   Patent: US 5912411-A 13 15-JUN-1999;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 9
AR095987/c
LOCUS       AR095987                38 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 13 from patent US 6004941.
ACCESSION   AR095987
VERSION     AR095987.1  GI:10024380
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Methods for regulating gene expression
  JOURNAL   Patent: US 6004941-A 13 21-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 10
AR136647/c
LOCUS       AR136647                38 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 13 from patent US 6136954.
ACCESSION   AR136647
VERSION     AR136647.1  GI:14477319
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Tetracycline-inducible transcriptional activator fusion proteins
  JOURNAL   Patent: US 6136954-A 13 24-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
```

```
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 11
AR157182/c
LOCUS       AR157182                  38 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 13 from patent US 6242667.
ACCESSION   AR157182
VERSION     AR157182.1  GI:15125886
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Transgenic organisms having tetracycline-regulated transcriptional
            regulatory systems
  JOURNAL   Patent: US 6242667-A 13 05-JUN-2001;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 12
AR164131/c
LOCUS       AR164131                  38 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 13 from patent US 6271348.
ACCESSION   AR164131
VERSION     AR164131.1  GI:16235083
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Tetracycline-inducible transcriptional inhibitor fusion proteins
  JOURNAL   Patent: US 6271348-A 13 07-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 13
I32817/c
LOCUS       I32817                    38 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 13 from patent US 5589362.
ACCESSION   I32817
VERSION     I32817.1  GI:1823608
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H., Gossen,M., Hillen,W., Helbl,V. and Schnappinger,D.
  TITLE     Tetracycline regulated transcriptional modulators with altered DNA
            binding specificities
  JOURNAL   Patent: US 5589362-A 13 31-DEC-1996;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 14
I59631/c
LOCUS       I59631                    38 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 13 from patent US 5654168.
ACCESSION   I59631
VERSION     I59631.1  GI:2478263
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Tetracycline-inducible transcriptional activator and
            tetracycline-regulated transcription units
  JOURNAL   Patent: US 5654168-A 13 05-AUG-1997;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 15
AX127254/c
LOCUS       AX127254                  38 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 13 from Patent EP1092771.
ACCESSION   AX127254
VERSION     AX127254.1  GI:14133334
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Bujard,H. and Gossen,M.
  TITLE     Tetracycline-regulated transcriptional modulators
  JOURNAL   Patent: EP 1092771-A 13 18-APR-2001;
            BASF AKTIENGESELLSCHAFT (DE) ; KNOLL AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 6;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 6.7e+05;
```

Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3
```

Search completed: May 26, 2004, 07:39:16
Job time : 1632.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:01:38 ; Search time 184.951 Seconds
(without alignments)
5030.293 Million cell updates/sec

Title: US-10-048-046-1_COPY_181_399
Perfect score: 219
Sequence: 1 gtcctcctgaggaagcggga.....tacagactggggatgtcatc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	20.2	9.2	50	4	AAL29957	Aal29957 Human SNP
	2	20	9.1	28	5	AAF30355	Aaf30355 Human che
	3	19.8	9.0	50	5	ABL01061	Abl01061 Human SNP
c	4	19.2	8.8	41	6	ABV76434	Abv76434 Human DNA
c	5	19	8.7	47	4	AAH88323	Aah88323 CNS disor
c	6	18.8	8.6	38	2	AAV35018	Aav35018 Human end
	7	18.8	8.6	47	3	AAZ67349	Aaz67349 Human map
c	8	18.4	8.4	38	2	AAT11362	Aat11362 Class C t
c	9	18.4	8.4	38	2	AAT45713	Aat45713 Class C T
c	10	18.4	8.4	38	2	AAV60085	Aav60085 Class C t
c	11	18.4	8.4	38	2	AAX60051	Aax60051 Class C t
c	12	18.4	8.4	38	3	AAZ56132	Aaz56132 Class C t
c	13	18.4	8.4	38	4	AAC61972	Aac61972 Nucleotid
c	14	18.4	8.4	38	4	AAH47637	Aah47637 Nucleotid
c	15	18.4	8.4	38	5	AAH25578	Aah25578 Nucleotid
c	16	18.4	8.4	38	6	ABK98648	Abk98648 Sequence
c	17	18.4	8.4	38	7	ACA94740	Aca94740 Class C T
c	18	18.4	8.4	38	7	ABX16493	Abx16493 Tet opera
c	19	18.4	8.4	38	8	ACD13899	Acd13899 Tetracycl
c	20	18.4	8.4	38	9	ADE52396	Ade52396 Tet opera
	21	18.4	8.4	46	2	AAT08997	Aat08997 Insulin r
	22	18.4	8.4	50	4	AAL28949	Aal28949 Human SNP
c	23	18.4	8.4	50	6	ABZ04239	Abz04239 Human leu
c	24	18.4	8.4	50	6	ABZ03712	Abz03712 Human leu
	25	18.2	8.3	33	9	ADD28164	Add28164 Reticuloc
c	26	18.2	8.3	41	6	AAL53837	Aal53837 Human cyc
c	27	18.2	8.3	43	7	ABZ21215	Abz21215 Nucleotid
	28	18.2	8.3	50	6	ABZ04919	Abz04919 Human leu
	29	18.2	8.3	50	6	ABZ06364	Abz06364 Human leu
	30	18.2	8.3	50	6	ABZ06168	Abz06168 Human leu
	31	18	8.2	29	2	AAX90453	Aax90453 Human sec
	32	18	8.2	29	4	AAS59360	Aas59360 Human sec
	33	18	8.2	29	6	ABA91029	Aba91029 Biotinyla
	34	18	8.2	42	2	AAV39800	Aav39800 Oligonucl
	35	18	8.2	50	3	AAA52575	Aaa52575 HCV RNA p
	36	18	8.2	50	3	AAA37201	Aaa37201 Human PRO
	37	18	8.2	50	4	AAL34067	Aal34067 Human SNP
	38	18	8.2	50	4	AAF54292	Aaf54292 Probe #22
	39	18	8.2	50	8	ACD68329	Acd68329 Novel hum
	40	18	8.2	50	8	ACH04431	Ach04431 Human sec
	41	18	8.2	50	8	ACD67975	Acd67975 Novel hum
	42	18	8.2	50	9	ADC17995	Adc17995 Human PRO
	43	18	8.2	50	9	ADD70641	Add70641 Human sec
	44	18	8.2	50	9	ADD39718	Add39718 Human sec
	45	18	8.2	50	9	ADD70164	Add70164 Human sec

ALIGNMENTS

RESULT 1
AAL29957/c
ID AAL29957 standard; DNA; 50 BP.
XX
AC AAL29957;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #3165.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Claim 1; Page 2293; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related

```
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-
CC   protein coupled receptors and thioesterases. The present sequence is one
CC   such oligonucleotide. The oligonucleotides and the peptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of the proteins listed above.
CC   Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms
XX
SQ   Sequence 50 BP; 8 A; 19 C; 11 G; 12 T; 0 U; 0 Other;

  Query Match             9.2%;  Score 20.2;  DB 4;  Length 50;
  Best Local Similarity   75.8%;  Pred. No. 6.9e+03;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        118 GGTCAGGTGACACTGGAAGATACCAGCACCAGT 150
              |||  ||||  ||||| | || |||||||| |
Db         39 GGTGGGGTGGAACTGGCACATGCCAGCACCACT 7


RESULT 2
AAF30355
ID   AAF30355 standard; DNA; 28 BP.
XX
AC   AAF30355;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Human checkpoint gene chfr 5' PCR primer.
XX
KW   Checkpoint with forkhead associated domain and ring finger; Chfr; human;
KW   mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;
KW   ubiquitin-protein ligase; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200109150-A2.
XX
PD   08-FEB-2001.
XX
PF   14-JUN-2000; 2000WO-US016391.
XX
PR   29-JUL-1999;   99US-0146194P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Halazonetis T,  Scolnick D;
XX
DR   WPI; 2001-182927/18.
XX
PT   Novel nucleic acid sequence of mitotic checkpoint gene encoding a
PT   checkpoint with forkhead-associated domain and ring finger protein, for
PT   diagnosing tumorigenic cells and in screening for anticancer drugs.
XX
PS   Example 3; Page 38; 85pp; English.
XX
CC   The present sequence is that of a 5' PCR primer, used with the 3' primer
CC   given in AAF30356, to amplify a cDNA fragment corresponding to
CC   nucleotides 352-1055 of the human chfr sequence given in AAF30352. The
CC   chfr gene encodes the human mitotic checkpoint protein Chfr (see
CC   AAB20219), which is required for regulation of the transition of cells
CC   from prophase to metaphase during mitosis. Loss of expression of Chfr is
CC   associated with a predisposition to tumourigenesis upon exposure to
CC   mitotic stress. A set of primers (see AAF30353-76) was used to amplify
CC   regions spanning the entire chfr coding region in order to determine
CC   whether the chfr gene is mutated in any of the human cancer cell lines
CC   SW480, DLD1, HT29, HCT116, SAOS2, U2OS, IMR5 and NGP. A mutation leading
CC   to a Val-580 to Met amino acid substitution was identified in the chfr
CC   gene of U2OS cells. Chfr polypeptides and chfr nucleic acids are used in
CC   methods of diagnosing tumourigenic cells and of screening for drugs which
CC   can inhibit the activity of Chfr in a cancer cell, rendering it more
CC   sensitive to additional antitumour therapies
XX
SQ   Sequence 28 BP; 12 A; 6 C; 6 G; 4 T; 0 U; 0 Other;

  Query Match             9.1%;  Score 20;  DB 5;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+03;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        130 CTGGAAGATACCAGCACCAG 149
              ||||||||||||||||||||
Db          9 CTGGAAGATACCAGCACCAG 28


RESULT 3
ABL01061
ID   ABL01061 standard; DNA; 50 BP.
XX
AC   ABL01061;
XX
DT   05-MAR-2002  (first entry)
XX
DE   Human SNP involving a gap oligonucleotide SEQ ID NO:1052.
XX
KW   Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW   immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW   autoimmune disease; inflammation; cancer; nervous system disease;
KW   infection; polymorphic protein; ds.
XX
OS   Homo sapiens.
XX
PN   WO200138586-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032311.
XX
PR   24-NOV-1999;   99US-0167383P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-355949/37.
XX
PT   Isolated human nucleic acids comprising one or more single nucleotide
PT   polymorphisms, useful for treating a subject suffering from a pathology,
PT   e.g. autoimmune diseases, ascribed to the presence of a sequence
PT   polymorphism.
XX
PS   Claim 1; Page 568; 674pp; English.
XX
CC   ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC   comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC   to ABB56903 represent human peptides encoded by some of the SNP
CC   oligonucleotides. The sequences from the present invention can have
CC   immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC   antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC   and antibodies from the present invention can be used for treating a
CC   subject suffering from, at risk for, or suspected of, suffering from a
CC   pathology ascribed to the presence of a sequence polymorphism. The
CC   pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC   the nervous system, and infection by pathogenic microorganisms. The SNPs
CC   are also useful for determining which forms of a characterised
CC   polymorphism are present in individuals. The antibodies may be used in
CC   the detection, quantitation and/or cellular or tissue localisation of a
CC   polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC   protein within appropriate physiological samples)
XX
SQ   Sequence 50 BP; 14 A; 3 C; 27 G; 6 T; 0 U; 0 Other;

  Query Match             9.0%;  Score 19.8;  DB 5;  Length 50;
```

```
  Best Local Similarity   69.2%;  Pred. No. 9.5e+03;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          9 GAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGTTG 47
              ||||| | ||  |||||  |   || ||||| |||| ||
Db          1 GAGGAGGAGGAGGTGGAGGAGGAGGGGGAGAAGAGGATG 39


RESULT 4
ABV76434/c
ID   ABV76434 standard; DNA; 41 BP.
XX
AC   ABV76434;
XX
DT   20-FEB-2003  (first entry)
XX
DE   Human DNA CGG repeat binding protein 10.23 probe, SEQ ID NO:9.
XX
KW   Human; DNA CGG repeat binding protein 10.23; recombinant production;
KW   gene therapy; embryonic development disorder; tumour; cancer; cytostatic;
KW   probe; ss.
XX
OS   Homo sapiens.
XX
PN   CN1343717-A.
XX
PD   10-APR-2002.
XX
PF   19-SEP-2000; 2000CN-00125256.
XX
PR   19-SEP-2000; 2000CN-00125256.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-548885/59.
XX
PT   A novel human DNA CGG-like repeated binder 10.23 polypeptide, useful for
PT   treating several diseases e.g. embryo development deformity and tumors.
XX
PS   Example 6; Page 22 (Disclosure); 35pp; Chinese.
XX
CC   The invention relates to human DNA CGG repeat binding protein 10.23
CC   (ABP58472) and nucleic acids encoding it (ABV76428). The protein has a
CC   molecular weight of 10.23 kD. The invention also relates to a method for
CC   the recombinant production of the protein, an antagonist of the protein,
CC   and the use of the protein, gene and antagonist in therapeutic
CC   applications. DNA CGG repeat binding protein 10.23 can be used in the
CC   treatment of a variety of diseases such as embryonic development
CC   disorders and tumours. Sequences ABV76433-ABV76434 represent human DNA
CC   CGG repeat binding protein 10.23 probes used in an exemplification of the
CC   invention
XX
SQ   Sequence 41 BP; 12 A; 11 C; 6 G; 12 T; 0 U; 0 Other;

  Query Match             8.8%;  Score 19.2;  DB 6;  Length 41;
  Best Local Similarity   67.5%;  Pred. No. 1.4e+04;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        149 GTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAGAAGCA 188
              ||| | || || |  | ||| || |  ||  |||||||||
Db         40 GTGTAGCATTGGTGTAAAAGGTGCACTTTCCTAAGAAGCA 1


RESULT 5
AAH88323/c
ID   AAH88323 standard; DNA; 47 BP.
XX
AC   AAH88323;
XX
DT   26-FEB-2002  (first entry)
XX
DE   CNS disorder-related biallelic marker #24 from CRFR1 gene.
XX
KW   Single nucleotide polymorphism; SNP; biallelic marker; human;
KW   central nervous system disorder; CNS; CRFR1;
KW   corticotrophin releasing factor receptor 1; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    24
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200151659-A2.
XX
PD   19-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-IB000116.
XX
PR   13-JAN-2000; 2000US-0175854P.
XX
PA   (GEST ) GENSET.
XX
PI   Chu T,  Blumenfeld M,  Cohen D;
XX
DR   WPI; 2001-483085/52.
XX
PT   Isolated polynucleotides, useful for genotyping nucleic acids for
PT   biallelic markers for the diagnosis of depression, comprises central
PT   nervous system disorder related biallelic marker.
XX
PS   Claim 1; Page 277; 519pp; English.
XX
CC   The present invention relates to biallelic markers derived from human
CC   genes involved in central nervous system (CNS) disorders. The present
CC   sequence is one such biallelic marker derived from human corticotrophin
CC   releasing factor receptor 1 (CRFR1) gene. CRF is a hormone released by
CC   the hypothalamus and stimulates the release of corticotrophin by the
CC   anterior pituitary gland. This marker has a single nucleotide
CC   polymorphism (SNP) and is useful in determining the genetic
CC   predisposition of individuals to CNS disorders, by identifying the
CC   nucleotides at a set of genetic markers in a biological sample, where the
CC   markers comprise at least one CNS disorder related marker
XX
SQ   Sequence 47 BP; 15 A; 8 C; 14 G; 9 T; 0 U; 1 Other;

  Query Match             8.7%;  Score 19;  DB 4;  Length 47;
  Best Local Similarity   67.6%;  Pred. No. 1.7e+04;
  Matches   25;  Conservative    1;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        170 TGAAGGTTGTTAAGAAGCAGACATGCCCTTTACAGAC 206
              || ||| |||| | :|||| | || ||| ||| ||
Db         38 TGTAGGATGTTTAAMAGCATTCCTGGCCTCTACCCAC 2


RESULT 6
AAV35018/c
ID   AAV35018 standard; DNA; 38 BP.
XX
AC   AAV35018;
XX
DT   27-AUG-1998  (first entry)
XX
DE   Human endothelin-beta 1 receptor PCR primer Vet122.
XX
KW   Endothelin beta-1 receptor; ETB-1; screening; signal transduction;
KW   agonist; antagonist; vasoconstrictor; vasopressor; cardiogenic shock;
KW   pulmonary hypertension; acute myocardial infarct; uraemia; vasculitis;
KW   Crohn's disease; ulcerative colitis; sepsis; congestive heart failure;
KW   coronary spasm; cyclosporin nephrotoxicity; toxaemia; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
```

```
XX
PN   US5773223-A.
XX
PD   30-JUN-1998.
XX
PF   02-SEP-1993;   93US-00117361.
XX
PR   02-SEP-1993;   93US-00117361.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Shyamala V,  Olson PT;
XX
DR   WPI; 1998-386992/33.
XX
PT   Screening for modulators of the endothelin B1 receptor - by measuring
PT   effect on signal transduction in cells engineered to express the
PT   receptor, potentially useful as agonists and antagonists of endothelin.
XX
PS   Example 2; Col 19; 23pp; English.
XX
CC   AAV35002-V35021 are primers used to amplify and isolate a novel human
CC   endothelin-beta1 receptor (ETB-1) which corresponds to a decapeptide
CC   insert. This sequence is used in a method involving the screening of
CC   compounds for their ability to bind to endothelin B1 (ETB1) receptor
CC   polypeptide and to modulate its signal transduction activity by applying
CC   test compound to host cells transformed with DNA encoding ETB-1 and
CC   optionally lysing the cells and then measuring signal transduction
CC   activity. The method is used to identify agonists and antagonists of
CC   endothelin (ET), a known vasoconstrictor/vasopressor agent, associated
CC   with cardiogenic shock, pulmonary hypertension, acute myocardial infarct,
CC   uraemia, Crohn's disease, ulcerative colitis, sepsis, congestive heart
CC   failure, coronary spasm, cyclosporin nephrotoxicity, vasculitis and
CC   toxaemia in pregnancy, and is also present at elevated levels after
CC   orthotopic liver transplantation and major abdominal surgery
XX
SQ   Sequence 38 BP; 9 A; 7 C; 5 G; 17 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 18.8;  DB 2;  Length 38;
  Best Local Similarity   68.4%;  Pred. No. 1.9e+04;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         65 GCAATAAACTGGTCTCTGGAGATCACTGTAGAATTGTA 102
              |||| |    | | |||| |||| |   ||||||||||
Db         38 GCAACAGCTCGATATCTGAAGATAAAAATAGAATTGTA 1


RESULT 7
AAZ67349
ID   AAZ67349 standard; DNA; 47 BP.
XX
AC   AAZ67349;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1696.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation; diagnosis;
KW   single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,G)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB000822.
XX
PR   21-APR-1998;   98US-0082614P.
PR   23-NOV-1998;   98US-0109732P.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome.
XX
PS   Claim 1; Page 592; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the invention
CC   have a variety of uses: they can be used for high density mapping of the
CC   human genome, and in complex association studies and haplotyping studies
CC   which are useful in determining the genetic basis for disease states.
CC   Compositions and methods of the invention can also be useful for the
CC   identification of the targets for the development of pharmaceutical
CC   agents and diagnostic methods, as well as the characterisation of the
CC   differential efficacious responses to and side effects from
CC   pharmaceutical agents acting on a disease as well as other treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC   3367, are not actually given a sequence in the Sequence Listing from the
CC   present invention
XX
SQ   Sequence 47 BP; 19 A; 8 C; 11 G; 9 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 18.8;  DB 3;  Length 47;
  Best Local Similarity   63.0%;  Pred. No. 2e+04;
  Matches   29;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy        140 CCAGCACCAGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAGAA 185
              | |||| | |||| ||   || |||||  ||||||  | |  | |
Db          1 CTAGCAGAACTGGATCATGAATCAACAAAATGAAGGAGGCTTTGCA 46


RESULT 8
AAT11362/c
ID   AAT11362 standard; DNA; 38 BP.
XX
AC   AAT11362;
XX
DT   07-JUL-1996  (first entry)
XX
DE   Class C tet operator sequence.
XX
KW   Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
KW   fusion protein; gene expression; regulation; inhibition; activation;
KW   transcription; ds.
XX
OS   Synthetic.
XX
PN   WO9601313-A1.
XX
PD   18-JAN-1996.
XX
PF   29-JUN-1995;   95WO-US008179.
XX
PR   01-JUL-1994;   94US-00270637.
PR   15-JUL-1994;   94US-00275876.
PR   03-FEB-1995;   95US-00383754.
PR   07-JUN-1995;   95US-00486814.
XX
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
XX
```

PI Bujard H, Gossen M;
XX
DR WPI; 1996-087666/09.
XX
PT New tetracycline-regulated transcription modulators - comprising fusion
PT proteins which bind to tet operator sequences to activate or inhibit
PT transcription.
XX
PS Disclosure; Page 76; 112pp; English.
XX
CC Fusion proteins comprising a first polypeptide which binds to a tet
CC operator sequence in the presence of tetracycline or a tetracycline
CC analogue, operatively linked to a second polypeptide which either
CC activates or inhibits transcription in eukaryotic cells may be used to
CC activate or inhibit transcription. Such proteins may be used to regulate
CC gene expression in cells and may be particularly useful for gene therapy
CC and for expression of gene products in transgenic organisms. Induction of
CC gene expression is rapid, efficient and strong, typically 1000-2000 fold.
CC The inducing agent does not cause pleitropic effects or cytotoxicity in
CC eukaryotic cells. The sequences of the different classes of tet operator
CC sequences are described in AAT11360-64)
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 2; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3
```

RESULT 9
AAT45713/c
ID AAT45713 standard; DNA; 38 BP.
XX
AC AAT45713;
XX
DT 18-MAR-1997 (first entry)
XX
DE Class C Tet operator sequence.
XX
KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
KW repressor; gene expression; therapy; transgenic animal; disease model;
KW ss.
XX
OS Escherichia coli.
XX
PN WO9640892-A1.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009049.
XX
PR 07-JUN-1995; 95US-00485971.
XX
PA (BADI) BASF AG.
XX
PI Bujard H, Gossen M, Hillen W, Helbl V, Schnappinger D;
XX
DR WPI; 1997-052305/05.
XX
PT Nucleic acid encoding tetracycline-inducible transcription regulatory
PT fusion protein - comprising modified tetracycline repressor able to bind
PT mutant tet operator, fused to transcription regulator, useful for
PT modulating eukaryotic gene expression.
XX
PS Disclosure; Page 82; 117pp; English.
XX
CC AAT45712-T45715 are class A, C, D and E tet operator sequences,
CC respectively. The main invention of the specification concerns modified
CC Tet repressor (TetR) proteins that bind to modified class B tet operator,
CC but also any other class of tet operator may be used. Modified class B
CC tet operators were modified at either posn. +4 or +6 by cytosine
CC substitutions to give tetO-4C and tetO6-C (see AAT45711 and AAT48478).
CC The tet operator sequences were used as targets for modified versions of
CC the Tet repressor (TetR). Modified TetR proteins can be fused to a
CC transcription regulatory polypeptide and used to control transcription of
CC a tetO-4C or tetO-6C linked gene. Nucleic acid encoding a fusion protein,
CC as above, may be introduced into a cell and transcription of the protein
CC can be controlled by altering the concn. of tetracycline (or an analogue)
CC in the cell. This ability to modulate gene expression in a predictable
CC way is very useful in gene therapy and for recombinant protein prodn. in
CC cultured cells or transgenic animals. The tetracycline (Tc)-inducible
CC system is also useful for the prodn. of transgenic animal models for the
CC study of disease and also for the study of gene function e.g. during
CC differentiation. The Tc-inducible system allows rapid activation of gene
CC transcription without cellular toxicity, high concns. of inducer are not
CC required
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 2; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3
```

RESULT 10
AAV60085/c
ID AAV60085 standard; DNA; 38 BP.
XX
AC AAV60085;
XX
DT 25-MAR-2003 (revised)
DT 04-DEC-1998 (first entry)
XX
DE Class C tet operator sequence.
XX
KW Tet repressor; tetracycline; regulation; expression;
KW Tet operator-linked gene; tet operator; ds.
XX
OS Synthetic.
XX
PN US5814618-A.
XX
PD 29-SEP-1998.
XX
PF 07-JUN-1995; 95US-00485978.
XX
PR 14-JUN-1993; 93US-00076327.
PR 14-JUN-1993; 93US-00076726.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
XX
PA (KNOL) KNOLL AG.
PA (BADI) BASF AG.
XX
PI Gossen M, Bujard H;
XX
DR WPI; 1998-541795/46.
XX
PT Tetracycline based regulation of gene expression - uses a tetracycline
PT operator sequence joined to a gene of interest, the gene of interest
PT being induced in the presence, but not absence of the antibiotic.
XX
PS Disclosure; Fig 5; 63pp; English.
XX
CC AAV60083-87 represent tet operator sequences of different classes. They
CC are used in the course of the invention. The specification describes a

CC method for regulating expression of a Tet (tetracycline) operator-linked
CC gene in a cell of a subject. The method comprises introducing into the
CC cell a nucleic acid encoding a fusion protein which inhibits
CC transcription in eukaryotic cells, the fusion protein comprising a
CC polypeptide which binds to a Tet operator sequence, operatively linked to
CC heterologous second polypeptide which inhibits transcription in
CC eukaryotic cells and modulating the concentration of a tetracycline
CC (analogue) in the subject. The method is used for the regulation of gene
CC expression system, using tetracycline (analogues). The system enables a
CC gene coupled to the system to be induced in the presence of Tet and then
CC stopped when Tet is removed. (Updated on 25-MAR-2003 to correct PR
CC field.)
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 2; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
| | ||| ||||| ||| |||| | | || ||||
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 11
AAX60051/c
ID AAX60051 standard; DNA; 38 BP.
XX
AC AAX60051;
XX
DT 20-MAR-2003 (revised)
DT 04-AUG-1999 (first entry)
XX
DE Class C tet operator sequence.
XX
KW Transgenic mice; transgene; tet operator-linked gene; tetracycline;
KW mouse-active transcriptional regulatory element; mutant Tet repressor;
KW gene therapy; genetic disease; acquired disease; cancer; viral disease;
KW vaccination; rheumatoid arthritis; hypopituitarism; wound healing;
KW tissue regeneration; cancer; benign prostatic hypertrophy; hemophilia;
KW erythrocytopenia; artherosclerosis; liver disease; Alzheimer's disease;
KW Parkinson's disease; human disease model; ds.
XX
OS Unidentified.
XX
PN US5912411-A.
XX
PD 15-JUN-1999.
XX
PF 07-JUN-1995; 95US-00487472.
XX
PR 14-JUN-1993; 93US-00076327.
PR 14-JUN-1993; 93US-00076726.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
XX
PA (UYHE-) UNIV HEIDELBERG.
XX
PI Bujard H, Gossen M;
XX
DR WPI; 1999-357232/30.
XX
PT Transgenic mice with inducible transgene activity useful for in vitro and
PT in vivo protein production.
XX
PS Disclosure; Fig 5; 63pp; English.
XX
CC The specification describes transgenic mice which have a transgene and a
CC tet operator-linked gene integrated in the genome. The transgene
CC comprises a mouse-active transcriptional regulatory element linked to a
CC polynucleotide sequence that encodes a fusion protein which activates
CC transcription of the tet operator-linked gene. The fusion protein
CC comprises a mutated Tet repressor that binds a tet operator sequence in
CC the presence of tetracycline (Tc) or a Tc analogue, linked to a
CC polypeptide that activates transcription in eukaryotic cells. The
CC transgenic system may be used for gene therapy to treat genes involved in
CC genetic or acquired diseases. Gene therapy may be used to treat cancer,
CC viral diseases, for vaccination, and to provide (Tc induced) regulated
CC doses of a product (e.g. for the treatment or regulation of rheumatoid
CC arthritis, hypopituitarism, wound healing and tissue regeneration,
CC cancer, benign prostatic hypertrophy, hemophilia, erythrocytopenia,
CC artherosclerosis and liver disease, Alzheimer's disease, and Parkinson's
CC disease). The system may also be used to produce proteins in vivo (e.g.
CC using mammalian, yeast or fungal cells) or in vitro (e.g. transgenic farm
CC animals), to produce animal models of human disease, or to produce a
CC stable cell line for gene cloning. The present sequence is used to
CC construct the transgenes of the invention. (Updated on 20-MAR-2003 to
CC correct PF field.)
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 2; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
| | ||| ||||| ||| |||| | | || ||||
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 12
AAZ56132/c
ID AAZ56132 standard; DNA; 38 BP.
XX
AC AAZ56132;
XX
DT 27-MAR-2000 (first entry)
XX
DE Class C tetracycline operator sequence.
XX
KW Class C tetracycline operator; gene expression regulation; cancer;
KW gene therapy; arthritis; wound healing; tissue regeneration; ds.
XX
OS Escherichia coli.
XX
PN US6004941-A.
XX
PD 21-DEC-1999.
XX
PF 07-JUN-1995; 95US-00485740.
XX
PR 14-JUN-1993; 93US-00076327.
PR 14-JUN-1993; 93US-00076726.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
XX
PA (BADI) BASF AG.
PA (BADI) BASF BIORESEARCH CORP.
PA (KNOL) KNOLL AG.
XX
PI Gossen M, Bujard H;
XX
DR WPI; 2000-085798/07.
XX
PT Regulation of gene expression in cells, useful for gene therapy of
PT diseases, production of proteins in vitro and in vivo and production of
PT stable cell lines for cloning.
XX
PS Disclosure; Fig 5; 64pp; English.
XX
CC This is a class C tetracycline operator sequence. The operator sequence
CC can be used in a method for the regulation of gene expression in a cell,

CC using tetracycline-regulated fusion proteins. The method involves
CC obtaining a cell from a subject, and introducing a nucleic acid molecule
CC into the cell, which operatively links a gene to at least one
CC tetracycline (tet) operator sequence. A second nucleic acid molecule is
CC then introduced which encodes a fusion protein, where the fusion protein
CC comprises a first polypeptide which binds to a tet operator sequence in
CC the presence of tetracycline, or its analogue, operatively linked to a
CC second polypeptide which activates transcription in eukaryotic cells to
CC form a modified cell. The modified cell can then be administered to the
CC subject, and the concentration of tetracycline (or and analogue) can be
CC regulated so that the expression of the gene is regulated. The method is
CC useful for gene therapy of diseases such as cancer and arthritis or for
CC tissue regeneration and wound healing. The method may also be useful for
CC the production of proteins in vitro and in vivo and for the production of
CC stable cell lines for cloning
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 3; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
| | ||| ||||| ||| |||| | | || ||||
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 13
AAC61972/c
ID AAC61972 standard; DNA; 38 BP.
XX
AC AAC61972;
XX
DT 06-MAR-2001 (first entry)
XX
DE Nucleotide sequence of a class C tet operator.
XX
KW Tetracycline; gene therapy; cancer; viral disease; rheumatoid arthritis;
KW hypopituitarism; wound healing; angiogenesis inhibitor; transgene;
KW immune response; tet operator; ds.
XX
OS Unidentified.
XX
PN US6136954-A.
XX
PD 24-OCT-2000.
XX
PF 28-SEP-1998; 98US-00162184.
XX
PR 14-JUN-1993; 93US-00076327.
PR 14-JUN-1993; 93US-00076726.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
PR 07-JUN-1995; 95US-00485978.
XX
PA (KNOL) KNOLL AG.
PA (BADI) BASF AG.
XX
PI Gossen M, Bujard H;
XX
DR WPI; 2001-040240/05.
XX
PT New tetracycline-inducible transcriptional activator fusion proteins
PT useful for regulating the level of gene expression or in gene therapy for
PT treating genetic or acquired diseases, e.g. as cancer, viral diseases, or
PT wound healing.
XX
PS Disclosure; Fig 5; 68pp; English.
XX
CC The specification describes methods for regulating gene expression in
CC subjects using tetracycline-responsive fusion protein. The fusion protein
CC comprises a polypeptide that binds to a test operator sequence in the
CC presence of tetracycline or tetracycline analogues, operatively linked to
CC a second polypeptide which activates transcription in eukaryotic cells.
CC The methods are used to turn gene expression on and off, or regulate the
CC level of gene expression. The system may be used in the study of cellular
CC development and differentiation in eukaryotic cells, plants and animals,
CC to regulate expression of site-specific recombinases (e.g. CRE or FLP) to
CC allow for irreversible modification of the genotype of a transgenic
CC organism under controlled conditions at a particular stage of
CC development. The system may further be used in gene therapy, in treating
CC either genetic or acquired diseases, such as cancer, viral diseases,
CC rheumatoid arthritis, hypopituitarism, or wound healing, and to
CC conditionally express suicide gene in cells. The regulatory system is
CC also used to express angiogenesis inhibitors from within a tumour via a
CC transgene regulated by this system, and avoid or inhibit an immune
CC response in subjects receiving treatments. The present sequence
CC represents a tet operator. It is used in the course of the invention
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 4; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
| | ||| ||||| ||| |||| | | || ||||
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 14
AAH47637/c
ID AAH47637 standard; DNA; 38 BP.
XX
AC AAH47637;
XX
DT 30-NOV-2001 (first entry)
XX
DE Nucleotide sequence of a class C tet operator.
XX
KW Fusion protein; tet operator; tetracycline; gene therapy; anti-cancer;
KW rheumatoid arthritis; hypopituitarism; wound healing; hemophilia;
KW diabetes; Alzheimer's disease; tet repressor; ds.
XX
OS Unidentified.
XX
PN US6271348-B1.
XX
PD 07-AUG-2001.
XX
PF 24-JAN-2000; 2000US-00489777.
XX
PR 14-JUN-1993; 93US-00076726.
PR 19-JUN-1993; 93US-00076327.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
PR 07-JUN-1995; 95US-00485978.
PR 28-SEP-1998; 98US-00162184.
XX
PA (BADI) BASF AG.
PA (KNOL) KNOLL AG.
XX
PI Bujard H, Gossen M;
XX
DR WPI; 2001-556625/62.
XX
PT Fusion protein for inhibiting transcription in eukaryotic cells useful in
PT gene therapy applications comprises a first polypeptide, which binds to
PT tet operator sequences, operatively linked to a heterologous second
PT polypeptide.
XX
PS Disclosure; Fig 5; 69pp; English.

XX
CC The invention relates to a fusion protein that comprises a first
CC polypeptide which binds to tet operator sequences, operatively linked to
CC a heterologous second polypeptide, which inhibits transcription in
CC eukaryotic cells. The fusion proteins are tetracycline-responsive and are
CC useful for regulation of transcription in eukaryotic cells and animals.
CC The tetracycline (Tc)-controlled regulatory system is useful in various
CC applications in gene therapy, such as in the treatment of various disease
CC conditions e.g. rheumatoid arthritis, hypopituitarism, wound healing and
CC tissue regeneration, anticancer treatments, benign prostatic hypertrophy,
CC hemophilia, diabretes and artherosclerosis. They are also useful for bone
CC marrow support therapy, treatment of central nervous system disorders
CC e.g. Alzheimer's disease, Parkinson's disease (see AAH47628 for a
CC detailed description of the uses). The present sequence represents the
CC nucleotide sequence of a class C tet operator
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 4; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
| | ||| ||||| ||| |||| | | || ||||
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 15
AAH25578/c
ID AAH25578 standard; DNA; 38 BP.
XX
AC AAH25578;
XX
DT 05-SEP-2001 (first entry)
XX
DE Nucleotide sequence of a tet operator, class A.
XX
KW Transgenic plant; transgene; tet operator-linked gene; Tet repressor;
KW tetracycline-regulated transcriptional regulatory system; tet operator;
KW tetracycline; ds.
XX
OS Unidentified.
XX
PN US6242667-B1.
XX
PD 05-JUN-2001.
XX
PF 28-SEP-1998; 98US-00161902.
XX
PR 14-JUN-1993; 93US-00076327.
PR 14-JUN-1993; 93US-00076726.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
PR 07-JUN-1995; 95US-00487472.
XX
PA (BADI) BASF AG.
PA (KNOL) KNOLL AG.
XX
PI Bujard H, Gossen M;
XX
DR WPI; 2001-396837/42.
XX
PT Transgenic plant for regulating the expression of genes, comprises a
PT transgene integrated into the genome and a tet operator-linked gene in
PT the genome.
XX
PS Disclosure; Fig 5; 65pp; English.
XX
CC The specification describes a transgenic plant with a tetracycline-
CC regulated transcriptional regulatory system. The transgenic plants have a
CC transgene and a tet operator-linked gene integrated in the genome, which
CC confers a detectable and functional phenotype on the plant when
CC expressed. The transgene comprises a transcriptional regulatory element
CC functional in cells of the plant operatively linked to a polynucleotide
CC sequence encoding a fusion protein that activates transcription of the
CC tet operator linked gene. The fusion protein comprises a first
CC polypeptide that is a mutated Tet repressor that binds to a tet operator
CC sequence in the presence of tetracycline or its analogue, operatively
CC linked to a second polypeptide that activates transcription in eukaryotic
CC cells. In the presence of tetracycline, the fusion protein binds to the
CC tet operator-linked gene and activates transcription of the tet operator-
CC linked gene such that it is expressed at a level sufficient to confer the
CC detectable and functional phenotype on the plant. The level of expression
CC of the tet-operator gene can be downmodulated by depleting tetracycline
CC from the plant. The transgene can be used to regulate the expression of
CC genes in the plant. The transgenic plant can be used to analyse the
CC functions of cellular proteins. The present sequence represents a tet
CC operator, which can be used to produce transgenic plants of the invention
XX
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 5; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.6e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
| | ||| ||||| ||| |||| | | || ||||
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

Search completed: May 26, 2004, 05:53:43
Job time : 188.951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:16:49 ; Search time 42.3039 Seconds
(without alignments)
2872.886 Million cell updates/sec

Title: US-10-048-046-1_COPY_181_399
Perfect score: 219
Sequence: 1 gtcctcctgaggaagcggga.....tacagactggggatgtcatc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	9.1	47	4	US-09-641-638-687	Sequence 687, App
c	2	18.8	8.6	38	1	US-08-117-361C-23	Sequence 23, Appl
c	3	18.4	8.4	38	1	US-08-485-971-13	Sequence 13, Appl
c	4	18.4	8.4	38	1	US-08-275-876-13	Sequence 13, Appl
c	5	18.4	8.4	38	1	US-08-383-754-13	Sequence 13, Appl
c	6	18.4	8.4	38	1	US-08-485-978-13	Sequence 13, Appl
c	7	18.4	8.4	38	2	US-08-486-814-13	Sequence 13, Appl
c	8	18.4	8.4	38	2	US-08-487-472-13	Sequence 13, Appl
c	9	18.4	8.4	38	3	US-08-485-740-13	Sequence 13, Appl
c	10	18.4	8.4	38	3	US-09-162-184-13	Sequence 13, Appl
c	11	18.4	8.4	38	3	US-09-161-902-13	Sequence 13, Appl
c	12	18.4	8.4	38	3	US-09-489-777A-13	Sequence 13, Appl
c	13	18.4	8.4	38	5	PCT-US95-08179-13	Sequence 13, Appl
	14	18.4	8.4	46	1	US-08-242-409-12	Sequence 12, Appl
	15	18.4	8.4	46	5	PCT-US95-05835-12	Sequence 12, Appl
	16	18.4	8.4	47	4	US-09-641-638-1120	Sequence 1120, Ap
	17	18.4	8.4	47	4	US-09-422-978-1696	Sequence 1696, Ap
c	18	17.8	8.1	49	1	US-08-171-389-54	Sequence 54, Appl
c	19	17.8	8.1	49	1	US-08-123-936-54	Sequence 54, Appl
c	20	17.8	8.1	49	2	US-08-475-228A-54	Sequence 54, Appl
c	21	17.8	8.1	49	3	US-08-482-080A-54	Sequence 54, Appl
c	22	17.8	8.1	49	4	US-09-354-947-54	Sequence 54, Appl
c	23	17.8	8.1	49	5	PCT-US93-12388-54	Sequence 54, Appl
	24	17.4	7.9	47	4	US-09-671-317-553	Sequence 553, App
	25	17.4	7.9	50	1	US-08-171-389-546	Sequence 546, App
	26	17.4	7.9	50	1	US-08-123-936-546	Sequence 546, App
	27	17.4	7.9	50	2	US-08-475-228A-546	Sequence 546, App
	28	17.4	7.9	50	3	US-08-482-080A-546	Sequence 546, App
	29	17.4	7.9	50	4	US-09-354-947-546	Sequence 546, App
	30	17.4	7.9	50	5	PCT-US93-12388-546	Sequence 546, App
	31	17.2	7.9	38	4	US-09-593-012-218	Sequence 218, App
c	32	17.2	7.9	41	1	US-07-931-473B-62	Sequence 62, Appl
c	33	17.2	7.9	41	1	US-07-714-131C-62	Sequence 62, Appl
c	34	17.2	7.9	41	1	US-08-412-110-62	Sequence 62, Appl
c	35	17.2	7.9	41	1	US-08-409-442A-62	Sequence 62, Appl
c	36	17.2	7.9	41	2	US-08-469-609A-62	Sequence 62, Appl
c	37	17.2	7.9	41	3	US-09-143-190-62	Sequence 62, Appl
c	38	17.2	7.9	41	4	US-09-502-344-62	Sequence 62, Appl
c	39	17.2	7.9	42	1	US-08-475-063-38	Sequence 38, Appl
c	40	17.2	7.9	42	1	US-08-207-792-38	Sequence 38, Appl
c	41	17.2	7.9	44	1	US-07-931-473B-106	Sequence 106, App
c	42	17.2	7.9	44	1	US-07-714-131C-106	Sequence 106, App
c	43	17.2	7.9	44	1	US-08-412-110-106	Sequence 106, App
c	44	17.2	7.9	44	1	US-08-409-442A-106	Sequence 106, App
c	45	17.2	7.9	44	2	US-08-469-609A-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-09-641-638-687
; Sequence 687, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 687
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 10-347-271 : polymorphic base A or T
US-09-641-638-687

Query Match 9.1%; Score 20; DB 4; Length 47;
Best Local Similarity 63.0%; Pred. No. 7.2e+02;
Matches 29; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

```
Qy         50 ACCTTTCCTTCCCCAGCAATAAACTGGTCTCTGGAGATCACTGTAG 95
              |||| ||||||||| ||     :||| |||||    | | ||| |
Db          2 ACCTCTCCTTCCCCTGCCTGGCWCTGTTCTCTCCTTAGTAGTGTGG 47
```

RESULT 2
US-08-117-361C-23/c
; Sequence 23, Application US/08117361C
; Patent No. 5773223
; GENERAL INFORMATION:
; APPLICANT: Venkatakrishna, Shyamala
; APPLICANT: Tekamp-Olson, Patricia

```
;    TITLE OF INVENTION:  Endothelin B1 (ETB1) Receptor Polypeptide
;    TITLE OF INVENTION:  Compositions, Methods, and Uses Thereof
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Streeet
;      CITY:  Emeryville
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  946087
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 &
;      SOFTWARE:  Word 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/117,361C
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Chung, Ling-Fong
;      REGISTRATION NUMBER:  36,482
;      REFERENCE/DOCKET NUMBER:  0945.001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 923-2704
;      TELEFAX:  (510) 655-3542
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-117-361C-23

  Query Match              8.6%;  Score 18.8;  DB 1;  Length 38;
  Best Local Similarity   68.4%;  Pred. No. 1.7e+03;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         65 GCAATAAACTGGTCTCTGGAGATCACTGTAGAATTGTA 102
              |||| |    | | |||| |||| |   ||||||||||
Db         38 GCAACAGCTCGATATCTGAAGATAAAAATAGAATTGTA 1


RESULT 3
US-08-485-971-13/c
; Sequence 13, Application US/08485971
; Patent No. 5589362
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    APPLICANT:  Hillen, Wolfgang
;    APPLICANT:  Helbl, Vera
;    APPLICANT:  Schnappinger, Dirk
;    TITLE OF INVENTION:  Tetracycline-Regulated Transcriptional
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,971
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/383,754
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP7
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-485-971-13

  Query Match              8.4%;  Score 18.4;  DB 1;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              |  |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 4
US-08-275-876-13/c
; Sequence 13, Application US/08275876
; Patent No. 5654168
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Tetracycline-Inducible Transcriptional
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/275,876
;      FILING DATE:
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US
;      FILING DATE:  01-JULY-94
```

```
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-275-876-13

  Query Match             8.4%;  Score 18.4;  DB 1;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 5
US-08-383-754-13/c
; Sequence 13, Application US/08383754
; Patent No. 5789156
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Tetracycline-Regulated Transcriptional
;    TITLE OF INVENTION:  Inhibitors
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/383,754
;      FILING DATE:
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-94
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-94
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-94
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-93
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
;      FILING DATE:  14-JUNE-93
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-383-754-13

  Query Match             8.4%;  Score 18.4;  DB 1;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 6
US-08-485-978-13/c
; Sequence 13, Application US/08485978
; Patent No. 5814618
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Methods for Regulating Gene Expression
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,978
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/383,754
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-485-978-13

  Query Match             8.4%;  Score 18.4;  DB 1;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 7
US-08-486-814-13/c
; Sequence 13, Application US/08486814
; Patent No. 5866755
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Animals Transgenic for a Tetracycline-    Controlled Transcr
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,814
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/383,754
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-486-814-13

  Query Match             8.4%;  Score 18.4;  DB 2;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 8
US-08-487-472-13/c
; Sequence 13, Application US/08487472
; Patent No. 5912411
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Animal Transgenic for a Tetracycline- Inducible Transcription
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,472
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/383,754
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-94
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
```

```
;       FILING DATE:  14-JUNE-1993
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DeConti, Giulio A. Jr.
;       REGISTRATION NUMBER:  31,503
;       REFERENCE/DOCKET NUMBER:  BBI-009CP3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617)227-7400
;       TELEFAX:  (617)227-5941
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  38 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-487-472-13

  Query Match             8.4%;  Score 18.4;  DB 2;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 9
US-08-485-740-13/c
; Sequence 13, Application US/08485740
; Patent No. 6004941
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Methods for Regulating Gene Expression
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,740
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/383,754
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-485-740-13

  Query Match             8.4%;  Score 18.4;  DB 3;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 10
US-09-162-184-13/c
; Sequence 13, Application US/09162184A
; Patent No. 6136954
;   GENERAL INFORMATION:
;       APPLICANT: Bujard, Hermann
;                  Gossen, Manfred
;       TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
;                           Activator Fusion Proteins
;       NUMBER OF SEQUENCES: 37
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/162,184A
;            FILING DATE: 28-Sep-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/383,754
;            FILING DATE: 03-FEB-1995
;            APPLICATION NUMBER: US 08/275,876
;            FILING DATE: 15-JULY-1994
;            APPLICATION NUMBER: US 08/270,637
;            FILING DATE: 01-JULY-1994
;            APPLICATION NUMBER: US 08/260,452
;            FILING DATE: 14-JUNE-1994
;            APPLICATION NUMBER: US 08/076,327
;            FILING DATE: 14-JUNE-1993
;            APPLICATION NUMBER: US 08/076,726
;            FILING DATE: 14-JUNE-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: DeConti, Giulio A. Jr.
;            REGISTRATION NUMBER: 31,503
;            REFERENCE/DOCKET NUMBER: BBI-009C6CN
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
```

```
;            TELEFAX: (617)742-4214
;   INFORMATION FOR SEQ ID NO: 13:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 38 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-162-184-13

  Query Match             8.4%;  Score 18.4;  DB 3;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 11
US-09-161-902-13/c
; Sequence 13, Application US/09161902
; Patent No. 6242667
;  GENERAL INFORMATION:
;    APPLICANT:  Bujard, Hermann
;    APPLICANT:  Gossen, Manfred
;    TITLE OF INVENTION:  Animal Transgenic for a Tetracycline- Inducible Transcription
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/161,902
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/487,472
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-1994
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-94
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,452
;      FILING DATE:  14-JUNE-1994
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,327
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/076,726
;      FILING DATE:  14-JUNE-1993
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009CP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-161-902-13

  Query Match             8.4%;  Score 18.4;  DB 3;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 12
US-09-489-777A-13/c
; Sequence 13, Application US/09489777A
; Patent No. 6271348
;  GENERAL INFORMATION:
;       APPLICANT: Bujard, Hermann
;                  Gossen, Manfred
;       TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
;                           Inhibitor Fusion Proteins
;       NUMBER OF SEQUENCES: 37
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 28 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/489,777A
;            FILING DATE: 24-Jan-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 09/162,184
;            FILING DATE: 28-SEP-1998
;            APPLICATION NUMBER: US 08/485,978
;            FILING DATE: 07-JUN-1995
;            APPLICATION NUMBER: US 08/383,754
;            FILING DATE: 03-FEB-1995
;            APPLICATION NUMBER: US 08/275,876
;            FILING DATE: 15-JULY-1994
;            APPLICATION NUMBER: US 08/270,637
;            FILING DATE: 01-JULY-1994
;            APPLICATION NUMBER: US 08/260,452
;            FILING DATE: 14-JUNE-1994
;            APPLICATION NUMBER: US 08/076,327
;            FILING DATE: 14-JUNE-1993
;            APPLICATION NUMBER: US 08/076,726
;            FILING DATE: 14-JUNE-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: DeConti, Giulio A. Jr.
;            REGISTRATION NUMBER: 31,503
;            REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)742-4214
;  INFORMATION FOR SEQ ID NO: 13:
;       SEQUENCE CHARACTERISTICS:
```

```
;             LENGTH: 38 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-489-777A-13

  Query Match              8.4%;  Score 18.4;  DB 3;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 13
PCT-US95-08179-13/c
; Sequence 13, Application PC/TUS9508179
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Tetracycline-Regulated Transcriptional
;    TITLE OF INVENTION:  Modulators
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/08179
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  To be assigned
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/383,754
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/275,876
;      FILING DATE:  15-JULY-94
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/270,637
;      FILING DATE:  01-JULY-94
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-009C2PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
PCT-US95-08179-13

  Query Match              8.4%;  Score 18.4;  DB 5;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.4e+03;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 14
US-08-242-409-12
; Sequence 12, Application US/08242409
; Patent No. 5496831
;  GENERAL INFORMATION:
;    APPLICANT:  Alexander-Bridges, Maria C.
;    APPLICANT:  Zhao, Hui-Fen
;    TITLE OF INVENTION:  INHIBITION OF INSULIN-INDUCED
;    TITLE OF INVENTION:  ADIPOSIS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/242,409
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/238001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-242-409-12

  Query Match              8.4%;  Score 18.4;  DB 1;  Length 46;
  Best Local Similarity   63.6%;  Pred. No. 2.7e+03;
  Matches   28;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         20 AGTGGACCATCGGGCGGAGACGAGGTTGCGACCTTTCCTTCCCC 63
              ||  ||   |  |||||||||   ||   |   ||||| ||||||
Db          2 AGGGGTTGAAAGGGCGGAGAGTCGGAAACTTTCTTTCTTTCCCC 45


RESULT 15
PCT-US95-05835-12
; Sequence 12, Application PC/TUS9505835
;  GENERAL INFORMATION:
;    APPLICANT:  Alexander-Bridges, Maria C.
;    APPLICANT:  Zhao, Hui-Fen
;    TITLE OF INVENTION:  INHIBITION OF INSULIN-
;    TITLE OF INVENTION:  INDUCED ADIPOSIS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/05835
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/242,409
;       FILING DATE:  13 May 1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00786/238001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  46
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
PCT-US95-05835-12

  Query Match             8.4%;  Score 18.4;  DB 5;  Length 46;
  Best Local Similarity   63.6%;  Pred. No. 2.7e+03;
  Matches   28;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         20 AGTGGACCATCGGGCGGAGACGAGGTTGCGACCTTTCCTTCCCC 63
              || ||   |  |||||||||   ||   |   ||||| ||||||
Db          2 AGGGGTTGAAAGGGCGGAGAGTCGGAAACTTTCTTTCTTTCCCC 45


Search completed: May 26, 2004, 09:20:40
Job time : 43.3039 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:41:55 ; Search time 955.965 Seconds
(without alignments)
1041.142 Million cell updates/sec

Title: US-10-048-046-1_COPY_181_399
Perfect score: 219
Sequence: 1 gtcctcctgaggaagcggga.....tacagactggggatgtcatc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 1612472

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	9.1	47	16	US-10-170-097-687	Sequence 687, App
	2	18.6	8.5	40	12	US-10-035-833A-3666	Sequence 3666, Ap
c	3	18.4	8.4	38	9	US-09-874-389-13	Sequence 13, Appl
c	4	18.4	8.4	38	10	US-09-921-650-13	Sequence 13, Appl
c	5	18.4	8.4	38	11	US-09-241-347-13	Sequence 13, Appl
c	6	18.4	8.4	38	15	US-10-032-393-64	Sequence 64, Appl
c	7	18.4	8.4	38	15	US-10-301-516-10	Sequence 10, Appl
c	8	18.4	8.4	38	15	US-10-326-671-53	Sequence 53, Appl
	9	18.4	8.4	47	16	US-10-170-097-1120	Sequence 1120, Ap
	10	18.4	8.4	47	16	US-10-349-143-1696	Sequence 1696, Ap
c	11	18.4	8.4	50	16	US-10-131-827-3703	Sequence 3703, Ap
c	12	18.4	8.4	50	16	US-10-131-827-4230	Sequence 4230, Ap
	13	18.2	8.3	46	13	US-10-252-155-739	Sequence 739, App
	14	18.2	8.3	46	13	US-10-252-155-740	Sequence 740, App
	15	18.2	8.3	50	16	US-10-131-827-4910	Sequence 4910, Ap
	16	18.2	8.3	50	16	US-10-131-827-6159	Sequence 6159, Ap
	17	18.2	8.3	50	16	US-10-131-827-6355	Sequence 6355, Ap
	18	18	8.2	29	9	US-09-729-674-247	Sequence 247, App
	19	18	8.2	42	9	US-09-179-536B-318	Sequence 318, App
	20	18	8.2	42	10	US-09-297-576A-318	Sequence 318, App
	21	18	8.2	50	10	US-09-946-374-126	Sequence 126, App
	22	18	8.2	50	12	US-10-015-395A-126	Sequence 126, App
	23	18	8.2	50	13	US-10-006-485A-126	Sequence 126, App
	24	18	8.2	50	13	US-10-013-907A-126	Sequence 126, App
	25	18	8.2	50	13	US-10-015-499A-126	Sequence 126, App
	26	18	8.2	50	13	US-10-226-254A-126	Sequence 126, App
	27	18	8.2	50	15	US-10-006-856A-126	Sequence 126, App
	28	18	8.2	50	15	US-10-006-818A-126	Sequence 126, App
	29	18	8.2	50	15	US-10-015-393A-126	Sequence 126, App
	30	18	8.2	50	15	US-10-015-869A-126	Sequence 126, App
	31	18	8.2	50	15	US-10-012-121A-126	Sequence 126, App
	32	18	8.2	50	15	US-10-006-116A-126	Sequence 126, App
	33	18	8.2	50	15	US-10-006-117A-126	Sequence 126, App
	34	18	8.2	50	15	US-10-017-527A-126	Sequence 126, App
	35	18	8.2	50	15	US-10-013-913A-126	Sequence 126, App
	36	18	8.2	50	15	US-10-007-194A-126	Sequence 126, App
	37	18	8.2	50	15	US-10-013-430A-126	Sequence 126, App
	38	18	8.2	50	15	US-10-011-671A-126	Sequence 126, App
	39	18	8.2	50	15	US-10-012-755A-126	Sequence 126, App
	40	18	8.2	50	15	US-10-015-386A-126	Sequence 126, App
	41	18	8.2	50	15	US-10-011-692A-126	Sequence 126, App
	42	18	8.2	50	15	US-10-006-768A-126	Sequence 126, App
	43	18	8.2	50	15	US-10-017-610A-126	Sequence 126, App
	44	18	8.2	50	15	US-10-006-063A-126	Sequence 126, App
	45	18	8.2	50	15	US-10-020-063A-126	Sequence 126, App

ALIGNMENTS

RESULT 1
US-10-170-097-687
; Sequence 687, Application US/10170097
; Publication No. US20030228582A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 687
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 10-347-271 : polymorphic base A or T
US-10-170-097-687

Query Match 9.1%; Score 20; DB 16; Length 47;

```
  Best Local Similarity   63.0%;  Pred. No. 8.5e+03;
  Matches   29;  Conservative    1;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         50 ACCTTTCCTTCCCCAGCAATAAACTGGTCTCTGGAGATCACTGTAG 95
              |||| |||||||||| ||     :||| |||||     | | ||| |
Db          2 ACCTCTCCTTCCCCTGCCTGGCWCTGTTCTCTCCTTAGTAGTGTGG 47


RESULT 2
US-10-035-833A-3666
; Sequence 3666, Application US/10035833A
; Publication No. US20040072156A1
; GENERAL INFORMATION:
;  APPLICANT: Nakamura, Yuho
;  APPLICANT:  Sekine, Akihiro
;  APPLICANT:  Iida, Aritoshi
;  APPLICANT:  Saito, Osamu
;  TITLE OF INVENTION: Detection of Genetic Polymorphisms
;  FILE REFERENCE: FORS-06904
;  CURRENT APPLICATION NUMBER: US/10/035,833A
;  CURRENT FILING DATE:  2001-12-27
;  NUMBER OF SEQ ID NOS: 7669
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3666
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (21)..(21)
;   OTHER INFORMATION: g is present or absent.
US-10-035-833A-3666

  Query Match             8.5%;  Score 18.6;  DB 12;  Length 40;
  Best Local Similarity   72.7%;  Pred. No. 2.4e+04;
  Matches   24;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 CTGAGGAAGCGGGAGTGGACCATCGGGCGGAGA 39
              | ||| ||| ||| |||||  || ||| || ||
Db          7 CAGAGTAAGGGGGGGTGGAGGATGGGGAGGGGA 39


RESULT 3
US-09-874-389-13/c
; Sequence 13, Application US/09874389
; Patent No. US20020152489A1
;   GENERAL INFORMATION:
;        APPLICANT: Bujard, Hermann
;                   Gossen, Manfred
;        TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible
;                            Transcription
;        NUMBER OF SEQUENCES: 28
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LAHIVE & COCKFIELD
;             STREET: 60 State Street, Suite 510
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02109-1875
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/874,389
;             FILING DATE: 26-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/161,902
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/275,876
;             FILING DATE: 15-JULY-1994
;             APPLICATION NUMBER: US 08/270,637
;             FILING DATE: 01-JULY-94
;             APPLICATION NUMBER: US 08/260,452
;             FILING DATE: 14-JUNE-1994
;             APPLICATION NUMBER: US 08/076,327
;             FILING DATE: 14-JUNE-1993
;             APPLICATION NUMBER: US 08/076,726
;             FILING DATE: 14-JUNE-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: DeConti, Giulio A. Jr.
;             REGISTRATION NUMBER: 31,503
;             REFERENCE/DOCKET NUMBER: BBI-009CP3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 13:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 38 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-874-389-13

  Query Match             8.4%;  Score 18.4;  DB 9;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.8e+04;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 4
US-09-921-650-13/c
; Sequence 13, Application US/09921650
; Publication No. US20030022315A1
;   GENERAL INFORMATION:
;        APPLICANT: Bujard, Hermann
;                   Gossen, Manfred
;        TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
;                            Activator Fusion Proteins
;        NUMBER OF SEQUENCES: 37
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LAHIVE & COCKFIELD
;             STREET: 28 State Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02109-1875
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/921,650
;             FILING DATE: 03-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 09/912,650
;             FILING DATE: 2001-08-03
;             APPLICATION NUMBER: US 08/485,978
;             FILING DATE: 07-JUN-1995
;             APPLICATION NUMBER: US 08/383,754
;             FILING DATE: 03-FEB-1995
;             APPLICATION NUMBER: US 08/275,876
;             FILING DATE: 15-JULY-1994
;             APPLICATION NUMBER: US 08/270,637
;             FILING DATE: 01-JULY-1994
;             APPLICATION NUMBER: US 08/260,452
;             FILING DATE: 14-JUNE-1994
```

```
;             APPLICATION NUMBER: US 08/076,327
;             FILING DATE: 14-JUNE-1993
;             APPLICATION NUMBER: US 08/076,726
;             FILING DATE: 14-JUNE-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: DeConti, Giulio A. Jr.
;             REGISTRATION NUMBER: 31,503
;             REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)742-4214
;   INFORMATION FOR SEQ ID NO: 13:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 38 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-921-650-13

  Query Match             8.4%;  Score 18.4;  DB 10;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.8e+04;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 5
US-09-241-347-13/c
; Sequence 13, Application US/09241347
; Publication No. US20040003417A1
;   GENERAL INFORMATION:
;        APPLICANT: Bujard, Hermann
;                   Gossen, Manfred
;        TITLE OF INVENTION: Animals Transgenic for a Tetracycline-    Controlled Transc
;        NUMBER OF SEQUENCES: 28
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LAHIVE & COCKFIELD
;             STREET: 60 State Street, Suite 510
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02109-1875
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/241,347
;             FILING DATE: 02-Feb-1999
;             CLASSIFICATION: 800
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/486,814
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/383,754
;             FILING DATE: 03-FEB-1995
;             APPLICATION NUMBER: US 08/275,876
;             FILING DATE: 15-JULY-1994
;             APPLICATION NUMBER: US 08/270,637
;             FILING DATE: 01-JULY-1994
;             APPLICATION NUMBER: US 08/260,452
;             FILING DATE: 14-JUNE-1994
;             APPLICATION NUMBER: US 08/076,327
;             FILING DATE: 14-JUNE-1993
;             APPLICATION NUMBER: US 08/076,726
;             FILING DATE: 14-JUNE-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: DeConti, Giulio A. Jr.
;             REGISTRATION NUMBER: 31,503
;             REFERENCE/DOCKET NUMBER: BBI-009CP4
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 13:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 38 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-241-347-13

  Query Match             8.4%;  Score 18.4;  DB 11;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.8e+04;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 6
US-10-032-393-64/c
; Sequence 64, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Gross, Molly
;  TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
;  FILE REFERENCE: ELITRA.010A
;  CURRENT APPLICATION NUMBER: US/10/032,393
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 60/259,434
;  PRIOR FILING DATE: 2000-12-27
;  PRIOR APPLICATION NUMBER: 09/948,993
;  PRIOR FILING DATE: 2001-09-06
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64
;   LENGTH: 38
;   TYPE: DNA
;   ORGANISM: Plasmid pSC101
US-10-032-393-64

  Query Match             8.4%;  Score 18.4;  DB 15;  Length 38;
  Best Local Similarity   69.4%;  Pred. No. 2.8e+04;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
              | |  ||| ||||| ||| |||| |  | || ||||
Db         38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3


RESULT 7
US-10-301-516-10/c
; Sequence 10, Application US/10301516
; Publication No. US20030180756A1
; GENERAL INFORMATION:
;  APPLICANT: SHI, YANG
;  APPLICANT:  SUI, GUANGCHAO
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SUPPRESSING EUKARYOTIC GENE
;  TITLE OF INVENTION:  EXPRESSION
;  FILE REFERENCE: HMV-084.01
;  CURRENT APPLICATION NUMBER: US/10/301,516
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: 60/366,478
;  PRIOR FILING DATE: 2002-03-21
;  NUMBER OF SEQ ID NOS: 39
```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic tet
; OTHER INFORMATION: operator sequence
US-10-301-516-10

Query Match 8.4%; Score 18.4; DB 15; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 8
US-10-326-671-53/c
; Sequence 53, Application US/10326671
; Publication No. US20030186281A1
; GENERAL INFORMATION:
; APPLICANT: Hillen, Wolfgang
; TITLE OF INVENTION: MODIFIED TETRACYCLINE REPRESSOR PROTEIN COMPOSITIONS AND METHODS
; TITLE OF INVENTION: USE
; FILE REFERENCE: 10182-022-999
; CURRENT APPLICATION NUMBER: US/10/326,671
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/343,278
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 459
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Plasmid pSC101
US-10-326-671-53

Query Match 8.4%; Score 18.4; DB 15; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAG 183
Db 38 AATTTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 9
US-10-170-097-1120
; Sequence 1120, Application US/10170097
; Publication No. US20030228582A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 1120
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 12-44-67 : polymorphic base T or C
US-10-170-097-1120

Query Match 8.4%; Score 18.4; DB 16; Length 47;
Best Local Similarity 73.3%; Pred. No. 3e+04;
Matches 22; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 63 CAGCAATAAACTGGTCTCTGGAGATCACTG 92
Db 14 CAGCCAGAAGYTGGTCTGTGGATGGCACTG 43

RESULT 10
US-10-349-143-1696
; Sequence 1696, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CP1
; CURRENT APPLICATION NUMBER: US/10/349,143
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US/09/422,978
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1696
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-5709-80 : polymorphic base A or G
US-10-349-143-1696

Query Match 8.4%; Score 18.4; DB 16; Length 47;
Best Local Similarity 60.9%; Pred. No. 3e+04;
Matches 28; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

Qy 140 CCAGCACCAGTGGAACAGTGATTAACAAGCTGAAGGTTGTTAAGAA 185
Db 1 CTAGCAGAACTGGATCATGAATCRACAAAATGAAGGAGGCTTTGCA 46

RESULT 11
US-10-131-827-3703/c
; Sequence 3703, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120

; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3703
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-3703

Query Match 8.4%; Score 18.4; DB 16; Length 50;
Best Local Similarity 63.6%; Pred. No. 3.1e+04;
Matches 28; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 37 AGACGAGGTTGCGACCTTTCCTTCCCCAGCAATAAACTGGTCTC 80
Db 47 AGACGAACTTGAGAACTTTCCTCCTCCTCCCTGGACCTCCTCCC 4

RESULT 12
US-10-131-827-4230/c
; Sequence 4230, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4230
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-4230

Query Match 8.4%; Score 18.4; DB 16; Length 50;
Best Local Similarity 69.4%; Pred. No. 3.1e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 50 ACCTTTCCTTCCCCAGCAATAAACTGGTCTCTGGAG 85
Db 46 ACATTTACATCCCCAGTAATGAAAGAGACTCAAGAG 11

RESULT 13
US-10-252-155-739
; Sequence 739, Application US/10252155
; Publication No. US20040068096A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT
; TITLE OF INVENTION: MULTI-DRUG RESISTANT PROTEINS
; FILE REFERENCE: D0152 NP
; CURRENT APPLICATION NUMBER: US/10/252,155
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/324,172
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/333,700
; PRIOR FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 783
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 739
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-252-155-739

Query Match 8.3%; Score 18.2; DB 13; Length 46;
Best Local Similarity 66.7%; Pred. No. 3.5e+04;
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 181 AAGAAGCAGACATGCCCTTTACAGACTGGGGATGTCATC 219
Db 2 AGGAAACAGCTATGACCTTGACATACATTGTGTTTCATC 40

RESULT 14
US-10-252-155-740
; Sequence 740, Application US/10252155
; Publication No. US20040068096A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT
; TITLE OF INVENTION: MULTI-DRUG RESISTANT PROTEINS
; FILE REFERENCE: D0152 NP
; CURRENT APPLICATION NUMBER: US/10/252,155
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/324,172
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/333,700
; PRIOR FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 783
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 740
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-252-155-740

Query Match 8.3%; Score 18.2; DB 13; Length 46;
Best Local Similarity 66.7%; Pred. No. 3.5e+04;
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 181 AAGAAGCAGACATGCCCTTTACAGACTGGGGATGTCATC 219
Db 2 AGGAAACAGCTATGACCTTGACATACATTGTGTTTCATC 40

RESULT 15
US-10-131-827-4910
; Sequence 4910, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4910
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-131-827-4910

Query Match 8.3%; Score 18.2; DB 16; Length 50;
Best Local Similarity 61.7%; Pred. No. 3.6e+04;
Matches 29; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy          91 TGTAGAATTGTAGTGGATGAAAAATCAGGTCAGGTGACACTGGAAGA 137
               || ||  || | |||  ||| | || || ||||| |      ||||
Db           2 TGGAGCGTTTTTGTGTTTGAGATATTAGCTCAGGTCAATTCCAAAGA 48
```

Search completed: May 26, 2004, 10:22:39
Job time : 955.965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:12:53 ; Search time 1520.36 Seconds
(without alignments)
4301.492 Million cell updates/sec

Title: US-10-048-046-1_COPY_181_399
Perfect score: 219
Sequence: 1 gtcctcctgaggaagcggga.....tacagactggggatgtcatc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		22	10.0	50	28	AZ481318	AZ481318 1M0303B05
2		21.6	9.9	50	9	AU102838	AU102838 AU102838
c 3		20.8	9.5	43	9	AI266869	AI266869 uj08a10.x
c 4		20.8	9.5	50	9	AU107911	AU107911 AU107911
c 5		20.2	9.2	50	9	AU105891	AU105891 AU105891
6		19.4	8.9	50	9	AU102843	AU102843 AU102843
7		19.4	8.9	50	9	AU102844	AU102844 AU102844
8		18.8	8.6	40	9	AA137304	AA137304 mq80b07.r
9		18.8	8.6	50	9	AL628983	AL628983 AL628983
10		18.8	8.6	50	9	AU102832	AU102832 AU102832
c 11		18.6	8.5	50	9	AU106276	AU106276 AU106276
c 12		18.6	8.5	50	9	AU106279	AU106279 AU106279
c 13		18.4	8.4	44	29	AL945796	AL945796 Arabidops
14		18.2	8.3	44	10	BE882963	BE882963 601505866
c 15		18.2	8.3	45	29	CG718854	CG718854 1119054G1
16		18.2	8.3	48	28	AZ486468	AZ486468 1M0314E04
c 17		18.2	8.3	50	9	AU104187	AU104187 AU104187
18		18	8.2	41	28	CC326813	CC326813 RRI150 Ba
c 19		18	8.2	43	28	BH906713	BH906713 SALK_0353
c 20		17.8	8.1	50	9	AU104910	AU104910 AU104910
21		17.6	8.0	40	9	AI634441	AI634441 ty98e04.x
c 22		17.6	8.0	46	29	CC887026	CC887026 SALK_1493
c 23		17.6	8.0	49	12	BJ065334	BJ065334 BJ065334
24		17.4	7.9	28	13	BQ593079	BQ593079 S015672-0
25		17.4	7.9	45	12	BI218941	BI218941 602938541
c 26		17.4	7.9	50	9	AU105889	AU105889 AU105889
c 27		17.2	7.9	33	28	AZ840354	AZ840354 2M0136C24
28		17.2	7.9	38	28	BZ292435	BZ292435 SALK_1242
29		17.2	7.9	38	29	TA117B04P	AL464666 T. brucei
30		17.2	7.9	41	28	AZ789628	AZ789628 2M0037A09
c 31		17.2	7.9	50	9	AU104911	AU104911 AU104911
c 32		17.2	7.9	50	9	AU106272	AU106272 AU106272
c 33		17.2	7.9	50	9	AU106275	AU106275 AU106275
c 34		17	7.8	43	28	BZ596218	BZ596218 SALK_0922
c 35		17	7.8	49	28	BZ380372	BZ380372 SALK_1150
36		17	7.8	50	9	AU102511	AU102511 AU102511
37		17	7.8	50	9	AU102514	AU102514 AU102514
38		17	7.8	50	9	AU102516	AU102516 AU102516
39		17	7.8	50	9	AU102517	AU102517 AU102517
c 40		17	7.8	50	9	AU106287	AU106287 AU106287
c 41		16.8	7.7	37	28	BH418190	BH418190 00-148-F1
c 42		16.8	7.7	49	28	BZ664689	BZ664689 SALK_1099
c 43		16.8	7.7	49	28	BZ664691	BZ664691 SALK_1099
c 44		16.8	7.7	50	9	AU106277	AU106277 AU106277
45		16.6	7.6	32	28	AZ332090	AZ332090 1M0060G18

ALIGNMENTS

RESULT 1
AZ481318
LOCUS AZ481318 50 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0303B05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0303B05 R, genomic survey sequence.
ACCESSION AZ481318
VERSION AZ481318.1 GI:10642383
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 50)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0303 row: B column: 05
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 50.

FEATURES Location/Qualifiers
source 1. .50
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0303B05"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 10.0%; Score 22; DB 28; Length 50;
Best Local Similarity 83.3%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        179 TTAAGAAGCAGACATGCCCTTTACAGACTG 208
              |||||| ||||||  | || ||||||||||
Db          4 TTAAGATGCAGACTGGTCCATTACAGACTG 33
```

RESULT 2
AU102838
LOCUS AU102838 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU102838 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone CAS09268, mRNA sequence.
ACCESSION AU102838
VERSION AU102838.1 GI:13552359
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CAS09268"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 9.9%; Score 21.6; DB 9; Length 50;
Best Local Similarity 75.0%; Pred. No. 3.7e+04;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy        107 ATGAAAAATCAGGTCAGGTGACACTGGAAGATACCA 142
              ||  ||||||||    || || ||||| ||||||||
Db          1 ATATAAAATCAGCGAGGGAGAAACTGGCAGATACCA 36
```

RESULT 3
AI266869/c
LOCUS AI266869 43 bp mRNA linear EST 16-NOV-1998
DEFINITION uj08a10.x1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1891290 3' similar to SW:ATP6_MOUSE P00848 ATP SYNTHASE A CHAIN ;, mRNA sequence.
ACCESSION AI266869
VERSION AI266869.1 GI:3885027
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 43)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:975614
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .43
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1891290"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor.[TGTTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3

vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Query Match 9.5%; Score 20.8; DB 9; Length 43;
Best Local Similarity 78.1%; Pred. No. 6e+04;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy         46 TGCGACCTTTCCTTCCCCAGCAATAAACTGGT 77
              || | |||||| ||||||| ||| |||  |||
Db         35 TGGGTCCTTTCTTTCCCCAACAAAAAATGGGT 4
```

RESULT 4
AU107911/c
LOCUS AU107911 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU107911 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone ZRV61770, mRNA sequence.
ACCESSION AU107911
VERSION AU107911.1 GI:13557433
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ZRV61770"
/clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

Query Match 9.5%; Score 20.8; DB 9; Length 50;
Best Local Similarity 64.6%; Pred. No. 6.5e+04;
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy          4 CTCCTGAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGTTGCGAC 51
              |  |  ||  |   |||||  |||||  |||||  |    ||  ||||| |||
Db         50 CACAGGACGGGACGGGACGGGACCGCCGGGCTGGCGCGGGGTTGGGAC 3
```

RESULT 5
AU105891/c
LOCUS AU105891 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU105891 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HTLB0018, mRNA sequence.
ACCESSION AU105891
VERSION AU105891.1 GI:13555412
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HTLB0018"
/clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

Query Match 9.2%; Score 20.2; DB 9; Length 50;
Best Local Similarity 65.1%; Pred. No. 9.9e+04;
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy          6 CCTGAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGTTGC 48
              || || |  | | | |  ||||| |||| | ||| ||||  ||
Db         48 CCAGAAGGCGAGCGNGGAGACCACCGGGGGAAGAGGAGGGNGC 6
```

RESULT 6
AU102843
LOCUS AU102843 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU102843 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone CAS11158, mRNA sequence.
ACCESSION AU102843
VERSION AU102843.1 GI:13552364
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS11158"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             8.9%;  Score 19.4;  DB 9;  Length 50;
  Best Local Similarity   70.3%;  Pred. No. 1.7e+05;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        106 GATGAAAAATCAGGTCAGGTGACACTGGAAGATACCA 142
              | |  |||| |||    || || ||||| ||||||||
Db          5 GCTATAAAAACAGCGAGGGAGAAACTGGCAGATACCA 41


RESULT 7
AU102844
LOCUS       AU102844                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU102844 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS11168, mRNA sequence.
ACCESSION   AU102844
VERSION     AU102844.1  GI:13552365
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS11168"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             8.9%;  Score 19.4;  DB 9;  Length 50;
  Best Local Similarity   70.3%;  Pred. No. 1.7e+05;
  Matches   26;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        106 GATGAAAAATCAGGTCAGGTGACACTGGAAGATACCA 142
              | |  |||| |||    || || ||||| ||||||||
Db          5 GCTATAAAAACAGCGAGGGAGAAACTGGCAGATACCA 41


RESULT 8
AA137304
LOCUS       AA137304                 40 bp    mRNA    linear   EST 09-FEB-1997
DEFINITION  mq80b07.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
            clone IMAGE:585013 5' similar to TR:G1002369 G1002369 COATOMER
            PROTEIN. ;, mRNA sequence.
ACCESSION   AA137304
VERSION     AA137304.1  GI:1700284
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:359661
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:585013"
                     /tissue_type="melanoma"
                     /dev_stage="M2 cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene mouse melanoma (#937312)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. From M2 cells, a highly metastatic derivative of
                     the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             8.6%;  Score 18.8;  DB 9;  Length 40;
  Best Local Similarity   76.7%;  Pred. No. 2.4e+05;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        171 GAAGGTTGTTAAGAAGCAGACATGCCCTTT 200
              | ||||| ||  ||| ||| ||||| ||||
Db          5 GCAGGTTCTTTTGAAACAGCCATGCGCTTT 34


RESULT 9
AL628983
LOCUS       AL628983                 50 bp    mRNA    linear   EST 19-NOV-2003
DEFINITION  AL628983 XGC-gastrula Silurana tropicalis cDNA clone TGas011g16 5',
            mRNA sequence.
ACCESSION   AL628983
VERSION     AL628983.1  GI:16598466
KEYWORDS    EST.
SOURCE      Silurana tropicalis (western clawed frog)
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Huckle E
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
```

```
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from stages 10-13
            gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas011g16.p1kSP6
            Sequencing primer: SP6.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Silurana tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas011g16"
                     /dev_stage="gastrula (stages 10.5-12 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /clone_lib="XGC-gastrula"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
ORIGIN

  Query Match             8.6%;  Score 18.8;  DB 9;  Length 50;
  Best Local Similarity   68.4%;  Pred. No. 2.7e+05;
  Matches   26;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          4 CTCCTGAGGAAGCGGGAGTGGACCATCGGGCGGAGACG 41
              ||||||   ||| |||||  ||||  |||  ||| | |
Db         13 CTCCTGTTCAAGGGGGAGAAGACCTGCGGCAGGACATG 50


RESULT 10
AU102832
LOCUS       AU102832                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU102832 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS01996, mRNA sequence.
ACCESSION   AU102832
VERSION     AU102832.1  GI:13552353
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS01996"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             8.6%;  Score 18.8;  DB 9;  Length 50;
  Best Local Similarity   76.7%;  Pred. No. 2.7e+05;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        113 AATCAGGTCAGGTGACACTGGAAGATACCA 142
              || |||    || |||||||| ||||||||
Db         10 AAACAGCGAGGGAGACACTGGCAGATACCA 39


RESULT 11
AU106276/c
LOCUS       AU106276                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106276 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC04067, mRNA sequence.
ACCESSION   AU106276
VERSION     AU106276.1  GI:13555797
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC04067"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             8.5%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   65.9%;  Pred. No. 3.1e+05;
  Matches   27;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          8 TGAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGTTGC 48
              || ||||  | | | |||  | ||| ||| | | |||| ||
Db         48 TGCGGAACAGCGCGCGGAGGAGCGGCCGGCGGCCAGGTCGC 8


RESULT 12
AU106279/c
LOCUS       AU106279                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106279 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC06165, mRNA sequence.
ACCESSION   AU106279
VERSION     AU106279.1  GI:13555800
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
```

```
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC06165"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             8.5%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   65.9%;  Pred. No. 3.1e+05;
  Matches   27;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          8 TGAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGTTGC 48
              || ||||  | | | |||  | ||| ||| | | |||| ||
Db         49 TGCGGAACAGCGCGCGGAGGAGCGGCCGGCGGCCAGGTCGC 9


RESULT 13
AL945796/c
LOCUS       AL945796                 44 bp    DNA     linear   GSS 24-OCT-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-291H10-015349,
            genomic survey sequence.
ACCESSION   AL945796
VERSION     AL945796.1  GI:24402418
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 44)
  AUTHORS   Li,Y., Strizhov,N., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At1g61850. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-291H10-015349"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             8.4%;  Score 18.4;  DB 29;  Length 44;
  Best Local Similarity   69.4%;  Pred. No. 3.3e+05;
  Matches   25;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        156 AGTGATTAACAAGCTGAAGGTTGTTAAGAAGCAGAC 191
              || ||| ||||    ||||| |  ||||||| || |
Db         43 AGCGATAAACATAAGGAAGGCTAATAAGAAGAAGGC 8


RESULT 14
BE882963
LOCUS       BE882963                 44 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601505866F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3907528 5',
            mRNA sequence.
ACCESSION   BE882963
VERSION     BE882963.1  GI:10331649
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9718  row: c  column: 17
            High quality sequence stop: 44.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3907528"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match             8.3%;  Score 18.2;  DB 10;  Length 44;
  Best Local Similarity   66.7%;  Pred. No. 3.8e+05;
  Matches   26;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          7 CTGAGGAAGCGGGAGTGGACCATCGGGCGGAGACGAGGT 45
              |||||| ||  || | ||||| | ||||    | ||  ||
Db          4 CTGAGGCAGAAGGTGGGGACCCTGGGGCCCTGCCGCTGT 42
```

```
RESULT 15
CG718854/c
LOCUS       CG718854                 45 bp    DNA     linear   GSS 20-OCT-2003
DEFINITION  1119054G10.1EL_y1 1119 - RescueMu Grid AA Zea mays genomic, genomic
            survey sequence.
ACCESSION   CG718854
VERSION     CG718854.1  GI:37750174
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1119054  row: 15
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73/K55"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="1119 - RescueMu Grid AA"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA
                     was extracted from leaf strips, double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
ORIGIN

  Query Match             8.3%;  Score 18.2;  DB 29;  Length 45;
  Best Local Similarity   66.7%;  Pred. No. 3.8e+05;
  Matches   26;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         44 GTTGCGACCTTTCCTTCCCCAGCAATAAACTGGTCTCTG 82
              | | |  ||| ||| ||   || ||| ||||||  ||||
Db         45 GATCCTCCCTCTCCATCAAGAGTAATGAACTGGGTTCTG 7


Search completed: May 26, 2004, 09:17:45
Job time : 1524.36 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:10:28 ; Search time 982.615 Seconds
(without alignments)
5822.506 Million cell updates/sec

Title: US-10-048-046-1_COPY_997_1128
Perfect score: 132
Sequence: 1 acatgcatcatctgccagga.....ctacctgccgctgtcccgtg 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20.2	15.3	47	6	AR290994	AR290994 Sequence
2		19.6	14.8	48	6	AX611835	AX611835 Sequence
c 3		19.2	14.5	47	6	I07024	I07024 Sequence 6
c 4		19.2	14.5	47	6	I27626	I27626 Sequence 18
5		19	14.4	47	6	I07021	I07021 Sequence 3
6		19	14.4	47	6	I27623	I27623 Sequence 15
c 7		18.2	13.8	36	6	AX740373	AX740373 Sequence
8		18.2	13.8	41	6	A01533	A01533 Joining reg
9		18.2	13.8	48	6	AR051077	AR051077 Sequence
10		18.2	13.8	48	6	AR430291	AR430291 Sequence
c 11		18	13.6	45	6	AR078468	AR078468 Sequence
c 12		18	13.6	47	6	A26491	A26491 Synthetic I
c 13		18	13.6	47	6	AR118511	AR118511 Sequence
c 14		18	13.6	47	6	AR153892	AR153892 Sequence
15		18	13.6	48	6	AX611836	AX611836 Sequence
16		17.6	13.3	42	6	E16421	E16421 DNA encodin
17		17.6	13.3	45	6	E16414	E16414 DNA encodin
18		17.6	13.3	48	6	E16420	E16420 DNA encodin
c 19		17.4	13.2	32	6	AR373963	AR373963 Sequence
20		17.4	13.2	40	6	AR363937	AR363937 Sequence
c 21		17.2	13.0	43	4	OCTR07	V00898 Rabbit trop
22		17.2	13.0	45	6	E16424	E16424 DNA encodin
23		17.2	13.0	45	6	E16427	E16427 DNA encodin
c 24		17.2	13.0	46	6	AR363127	AR363127 Sequence
c 25		17.2	13.0	46	6	AX769580	AX769580 Sequence
26		17	12.9	36	6	I91771	I91771 Sequence 5
c 27		17	12.9	47	6	AR289000	AR289000 Sequence
c 28		17	12.9	47	6	AX378744	AX378744 Sequence
29		17	12.9	48	6	AR031652	AR031652 Sequence
30		17	12.9	48	6	AR032616	AR032616 Sequence
31		17	12.9	48	6	I29356	I29356 Sequence 22
32		17	12.9	48	6	I90270	I90270 Sequence 11
33		17	12.9	48	6	I91030	I91030 Sequence 22
34		17	12.9	48	6	AR209280	AR209280 Sequence
c 35		16.8	12.7	30	6	I89979	I89979 Sequence 27
c 36		16.8	12.7	39	6	A22325	A22325 Primer O-2
37		16.8	12.7	39	6	A22326	A22326 Primer O-2
38		16.8	12.7	40	6	AR253252	AR253252 Sequence
39		16.8	12.7	40	6	BD138012	BD138012 Lectomedi
c 40		16.8	12.7	47	6	AR001266	AR001266 Sequence
c 41		16.8	12.7	47	6	I23424	I23424 Sequence 17
c 42		16.8	12.7	48	6	BD182587	BD182587 Anti TRAI
c 43		16.8	12.7	50	6	AR032939	AR032939 Sequence
44		16.8	12.7	50	6	E50843	E50843 Process for
c 45		16.8	12.7	50	6	I29679	I29679 Sequence 55

ALIGNMENTS

RESULT 1
AR290994
LOCUS AR290994 47 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 2729 from patent US 6537751.
ACCESSION AR290994
VERSION AR290994.1 GI:31678278
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 47)
AUTHORS Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE Biallelic markers for use in constructing a high density disequilibrium map of the human genome
JOURNAL Patent: US 6537751-A 2729 25-MAR-2003;

```
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             15.3%;  Score 20.2;  DB 6;  Length 47;
  Best Local Similarity   65.1%;  Pred. No. 1.4e+06;
  Matches   28;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         17 AGGACCTGCTGCACGACTGCGTGAGTTTGCAGCCCTGCATGCA 59
              ||| ||||| |||  |  |:  |||||| |  ||| ||| ||
Db          5 AGGTCCTGCAGCAGTATGGYACCAGTTTGAAAGCCTCCATCCA 47


RESULT 2
AX611835
LOCUS       AX611835                 48 bp    DNA     linear   PAT 17-FEB-2003
DEFINITION  Sequence 2860 from Patent WO02072882.
ACCESSION   AX611835
VERSION     AX611835.1  GI:28407264
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cullen,P. and Seedorf,U.
  TITLE     Coronary chip
  JOURNAL   Patent: WO 02072882-A 2860 19-SEP-2002;
            OGHAM GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             14.8%;  Score 19.6;  DB 6;  Length 48;
  Best Local Similarity   73.5%;  Pred. No. 1.9e+06;
  Matches   25;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         23 TGCTGCACGACTGCGTGAGTTTGCAGCCCTGCAT 56
              ||  |  |||||| |  ||| ||| ||||| ||||
Db         12 TGATAAACGACTTCAAGAGCTTGGAGCCCGGCAT 45


RESULT 3
I07024/c
LOCUS       I07024                   47 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 6 from Patent EP 0314317.
ACCESSION   I07024
VERSION     I07024.1  GI:590377
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Capon,D.J. and Gregory,T.J.
  TITLE     Adheson variants, nucleic acid encoding them and compositions
            comprising them
  JOURNAL   Patent: EP 0314317-A1 6 03-MAY-1989;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             14.5%;  Score 19.2;  DB 6;  Length 47;
  Best Local Similarity   67.5%;  Pred. No. 2.3e+06;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCC 113
              |||||| || | |||| |    |||||| |||| |   ||
Db         46 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCC 7


RESULT 4
I27626/c
LOCUS       I27626                   47 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 18 from patent US 5565335.
ACCESSION   I27626
VERSION     I27626.1  GI:1818402
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Capon,D.J. and Gregory,T.J.
  TITLE     Adhesion variants
  JOURNAL   Patent: US 5565335-A 18 15-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             14.5%;  Score 19.2;  DB 6;  Length 47;
  Best Local Similarity   67.5%;  Pred. No. 2.3e+06;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCC 113
              |||||| || | |||| |    |||||| |||| |   ||
Db         46 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCC 7


RESULT 5
I07021
LOCUS       I07021                   47 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 3 from Patent EP 0314317.
ACCESSION   I07021
VERSION     I07021.1  GI:590374
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Capon,D.J. and Gregory,T.J.
  TITLE     Adheson variants, nucleic acid encoding them and compositions
            comprising them
  JOURNAL   Patent: EP 0314317-A1 3 03-MAY-1989;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             14.4%;  Score 19;  DB 6;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 2.6e+06;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTG 108
              |||||| || | |||| |    |||||| |||| |
Db         12 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAG 46


RESULT 6
I27623
LOCUS       I27623                   47 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 15 from patent US 5565335.
ACCESSION   I27623
VERSION     I27623.1  GI:1818399
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Capon,D.J. and Gregory,T.J.
  TITLE     Adhesion variants
  JOURNAL   Patent: US 5565335-A 15 15-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           14.4%;  Score 19;  DB 6;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 2.6e+06;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTG 108
              |||||| || | |||| |    |||||| |||| |
Db         12 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAG 46


RESULT 7
AX740373/c
LOCUS       AX740373                 36 bp    DNA     linear   PAT 10-MAY-2003
DEFINITION  Sequence 107 from Patent EP1300419.
ACCESSION   AX740373
VERSION     AX740373.1  GI:30523546
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Buettner,C., Schwarz,M., Knackmuss,S., Peter,K., Roettgen,P. and
            Little,M.
  TITLE     Antibody of human origin for inhibiting thrombocyte aggregation
  JOURNAL   Patent: EP 1300419-A 107 09-APR-2003;
            Affimed Therapeutics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide"
ORIGIN

  Query Match           13.8%;  Score 18.2;  DB 6;  Length 36;
  Best Local Similarity   87.0%;  Pred. No. 4.3e+06;
  Matches   20;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         46 CAGCCCTGCATGCACACGTTCTG 68
              |||||||||  || ||  |||||||
Db         27 CAGCCCTGCCTGTACGCGTTCTG 5


RESULT 8
A01533
LOCUS       A01533                   41 bp    DNA     linear   PAT 25-FEB-1993
DEFINITION  Joining region of the bacterial origin and part of pBR322 origin in
            pBB9b28.
ACCESSION   A01533
VERSION     A01533.1  GI:344373
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   .
  TITLE     EXPRESSION VECTORS AND METHOD FOR THEIR CONSTRUCTION
  JOURNAL   Patent: WO 8902466-A 4 23-MAR-1989;
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           13.8%;  Score 18.2;  DB 6;  Length 41;
  Best Local Similarity   74.2%;  Pred. No. 4.2e+06;
  Matches   23;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         40 AGTTTGCAGCCCTGCATGCACACGTTCTGCG 70
              ||| |||| |||||| ||| | | ||| | |
Db         10 AGTATGCAACCCTGCCTGCGCGCTTTCGGTG 40


RESULT 9
AR051077
LOCUS       AR051077                 48 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 20 from patent US 5830650.
ACCESSION   AR051077
VERSION     AR051077.1  GI:5974441
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Crea,R.
  TITLE     Walk-through mutagenesis
  JOURNAL   Patent: US 5830650-A 20 03-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           13.8%;  Score 18.2;  DB 6;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 4e+06;
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    |  ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 10
AR430291
LOCUS       AR430291                 48 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 20 from patent US 6649340.
ACCESSION   AR430291
VERSION     AR430291.1  GI:40191060
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Crea,R.
  TITLE     Walk-through mutagenesis
  JOURNAL   Patent: US 6649340-A 20 18-NOV-2003;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           13.8%;  Score 18.2;  DB 6;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 4e+06;
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    |  ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 11
AR078468/c
LOCUS       AR078468                 45 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 8 from patent US 5962664.
ACCESSION   AR078468
```

```
VERSION     AR078468.1  GI:10005214
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Friedhoff,A.J., Basham,D.A. and Miller,J.C.
  TITLE     Psychosis protecting nucleic acid, peptides, compositions and
            method of use
  JOURNAL   Patent: US 5962664-A 8 05-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             13.6%;  Score 18;  DB 6;  Length 45;
  Best Local Similarity   80.8%;  Pred. No. 4.6e+06;
  Matches   21;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGC 99
              |||| | | | ||||||||||||| ||
Db         29 CTTGTTGCCCAGGCTGGATGGAGTGC 4


RESULT 12
A26491/c
LOCUS       A26491                     47 bp    DNA     linear   PAT 01-MAR-2003
DEFINITION  Synthetic IDUA probe ID47.
ACCESSION   A26491
VERSION     A26491.1  GI:1248227
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Hopwood,J.J., Orsborn,A.M., Anson,D.S., Clements,P.R., Morris,C.P.,
            Nelson,P.V. and Scott,H.S.
  TITLE     Synthetic alpha-L-iduronidase and genetic sequences encoding same
  JOURNAL   Patent: WO 9310244-A 1 27-MAY-1993;
            WOMENS & CHILDRENS HOSPITAL (AU)
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             13.6%;  Score 18;  DB 6;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 4.5e+06;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
              ||||| |    | | ||||  ||||||| | |  ||  ||||
Db         47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 13
AR118511/c
LOCUS       AR118511                   47 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6149909.
ACCESSION   AR118511
VERSION     AR118511.1  GI:14100421
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Scott,H.Steele., Anson,D.Stewart., Orsborn,A.Marie.,
            Nelson,P.Victor., Clements,P.Roy., Morris,C.Phillip. and
            Hopwood,J.Joseph.
  TITLE     Synthetic .alpha.-L-iduronidase and genetic sequences encoding same
  JOURNAL   Patent: US 6149909-A 1 21-NOV-2000;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             13.6%;  Score 18;  DB 6;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 4.5e+06;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
              ||||| |    | | ||||  ||||||| | |  ||  ||||
Db         47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 14
AR153892/c
LOCUS       AR153892                   47 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6238662.
ACCESSION   AR153892
VERSION     AR153892.1  GI:15121945
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Scott,H.Steel., Anson,D.Stewart., Orsborn,A.Marie.,
            Nelson,P.Victor., Clements,P.Roy., Morris,C.Phillip. and
            Hopwood,J.Joseph.
  TITLE     Synthetic .alpha.-L-iduronidase and genetic sequences encoding same
  JOURNAL   Patent: US 6238662-A 1 29-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             13.6%;  Score 18;  DB 6;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 4.5e+06;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
              ||||| |    | | ||||  ||||||| | |  ||  ||||
Db         47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 15
AX611836
LOCUS       AX611836                   48 bp    DNA     linear   PAT 17-FEB-2003
DEFINITION  Sequence 2861 from Patent WO02072882.
ACCESSION   AX611836
VERSION     AX611836.1  GI:28407265
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cullen,P. and Seedorf,U.
  TITLE     Coronary chip
  JOURNAL   Patent: WO 02072882-A 2861 19-SEP-2002;
            OGHAM GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             13.6%;  Score 18;  DB 6;  Length 48;
  Best Local Similarity   70.6%;  Pred. No. 4.5e+06;
  Matches   24;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy          23 TGCTGCACGACTGCGTGAGTTTGCAGCCCTGCAT 56
               ||  |  | |||| |  ||| ||| ||||| ||||
Db          12 TGATAAATGACTTCAAGAGCTTGGAGCCCGGCAT 45
```

Search completed: May 26, 2004, 07:39:18
Job time : 984.615 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:01:38 ; Search time 111.477 Seconds
(without alignments)
5030.293 Million cell updates/sec

Title: US-10-048-046-1_COPY_997_1128
Perfect score: 132
Sequence: 1 acatgcatcatctgccagga.....ctacctgccgctgtcccgtg 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20.2	15.3	42	6	ABL95873	Abl95873 LLPL-rela
c	2	19.2	14.5	50	9	ADD31850	Add31850 Synthetic
	3	19	14.4	47	3	AAZ68382	Aaz68382 Human map
	4	18.8	14.2	41	6	ABN83043	Abn83043 Human ank
c	5	18.2	13.8	36	9	ADE03377	Ade03377 Human imm
	6	18.2	13.8	48	2	AAV60897	Aav60897 Oligonucl
	7	18.2	13.8	48	9	ADD54968	Add54968 Heavy cha
c	8	18	13.6	47	2	AAQ42415	Aaq42415 IDUA prim
c	9	18	13.6	48	7	ABS56353	Abs56353 PCR prime
	10	18	13.6	50	6	ABZ01079	Abz01079 Human leu
c	11	17.8	13.5	50	4	AAL31778	Aal31778 Human SNP
	12	17.6	13.3	42	2	AAV02057	Aav02057 Mutated h
	13	17.6	13.3	45	2	AAV02046	Aav02046 hCS-15 pe
	14	17.6	13.3	48	2	AAV02056	Aav02056 Mutated h
	15	17.6	13.3	50	4	AAL34532	Aal34532 Human SNP
c	16	17.4	13.2	28	5	AAF30356	Aaf30356 Human che
	17	17.4	13.2	49	8	ACF04078	Acf04078 Cdc20 sup
	18	17.2	13.0	41	6	ABN83042	Abn83042 Human ank
c	19	17.2	13.0	41	6	ABV75713	Abv75713 Tumour su
	20	17.2	13.0	45	2	AAV02060	Aav02060 Mutated h
	21	17.2	13.0	45	2	AAV02063	Aav02063 Mutated h
c	22	17.2	13.0	46	1	AAN60796	Aan60796 Sequence
c	23	17.2	13.0	46	2	AAT28216	Aat28216 Feline he
c	24	17.2	13.0	46	7	ABZ68916	Abz68916 PCR prime
	25	17.2	13.0	50	4	AAL33784	Aal33784 Human SNP
	26	17	12.9	36	2	AAQ97839	Aaq97839 Human E1
	27	17	12.9	36	2	AAT41762	Aat41762 Ubiquitin
	28	17	12.9	36	2	AAX78503	Aax78503 Human E1
	29	17	12.9	36	3	AAA61629	Aaa61629 Human E1
c	30	17	12.9	47	3	AAZ66388	Aaz66388 Human map
c	31	17	12.9	47	6	ABK41285	Abk41285 Human LSR
	32	17	12.9	48	2	AAQ69478	Aaq69478 Human leu
	33	17	12.9	48	2	AAT63940	Aat63940 Human LFA
	34	17	12.9	48	2	AAX17228	Aax17228 Test sequ
	35	17	12.9	48	6	ABK82719	Abk82719 DNA bindi
	36	17	12.9	48	10	ADE80258	Ade80258 Duplex ol
c	37	17	12.9	50	2	AAV07319	Aav07319 Nucleotid
c	38	17	12.9	50	4	AAL33674	Aal33674 Human SNP
c	39	17	12.9	50	4	AAL30334	Aal30334 Human SNP
c	40	16.8	12.7	38	4	AAF54939	Aaf54939 PCR prime
	41	16.8	12.7	40	2	AAZ28006	Aaz28006 Human lec
	42	16.8	12.7	40	3	AAA96033	Aaa96033 Human lec
c	43	16.8	12.7	47	2	AAQ28991	Aaq28991 Sequence
	44	16.8	12.7	47	3	AAZ67208	Aaz67208 Human map
c	45	16.8	12.7	47	3	AAZ68974	Aaz68974 Human map

ALIGNMENTS

RESULT 1
ABL95873
ID ABL95873 standard; DNA; 42 BP.
XX
AC ABL95873;
XX
DT 19-JUN-2002 (first entry)
XX
DE LLPL-related oligonucleotide #2.
XX
KW Recombinant protein production; drug; reagent; food stuff; ss.
XX
OS Unidentified.
XX
PN WO200208417-A1.
XX
PD 31-JAN-2002.
XX
PF 25-JUL-2001; 2001WO-JP006392.
XX
PR 25-JUL-2000; 2000JP-00229064.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Ito T, Tanaka Y, Kondo M;
XX
DR WPI; 2002-179906/23.
XX
PT Production of recombinant proteins in prokaryotes or eukaryotes
PT particularly with target proteins obtainable through gene recombination
PT technique, for use as drugs, reagents, raw materials for industries and
PT feeding stuffs.
XX
PS Example 23; Fig 29; 137pp; Japanese.
XX
CC The present invention relates to a method for producing recombinant
CC proteins. The method comprises preparing a recombinant vector for
CC transforming a host cell before culturing the obtained transformant,
CC assaying expression of the reporter gene and confirming high expression
CC of the reporter gene. The recombinant proteins are useful as drugs,
CC reagents, raw materials for industries and feeding stuffs. Also, the
CC proteins are obtainable on large-scale production. The present sequence
CC was used to illustrate the invention
XX
SQ Sequence 42 BP; 3 A; 16 C; 16 G; 7 T; 0 U; 0 Other;

```
  Query Match           15.3%;  Score 20.2;  DB 6;  Length 42;
  Best Local Similarity  68.3%;  Pred. No. 1.6e+04;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         10 ATCTGCCAGGACCTGCTGCACGACTGCGTGAGTTTGCAGCC 50
              ||  ||| | |||||| || | || | ||| |  ||| |||
Db          1 ATGGGCCTGCACCTGCGGCCCTACCGGGTGGGCCTGCTGCC 41


RESULT 2
ADD31850/c
ID   ADD31850 standard; DNA; 50 BP.
XX
AC   ADD31850;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Synthetic antibody oligonucleotide SEQ ID NO:89.
XX
KW   recombination product; synthetic gene technology; antibody; ss.
XX
OS   Synthetic.
XX
PN   WO2003064611-A2.
XX
PD   07-AUG-2003.
XX
PF   29-JAN-2003; 2003WO-US002612.
XX
PR   30-JAN-2002; 2002US-00062188.
XX
PA   (EGEA-) EGEA BIOSCIENCES INC.
XX
PI   Evans GA;
XX
DR   WPI; 2003-663477/62.
XX
PT   Creating recombination products between two distinct nucleotide
PT   sequences, useful in the field of synthetic gene technology, and in
PT   assembling a library, or a population or a collection of polypeptide
PT   variants.
XX
PS   Example 2; SEQ ID NO 89; 132pp; English.
XX
CC   The present invention describes a method for creating a collection of
CC   recombination products between two nucleotide sequences. The method
CC   comprises combining an initial set of oligonucleotides corresponding to a
CC   first nucleotide sequence with a subsequent set of oligonucleotides
CC   corresponding to a distinct nucleotide sequence and further combining the
CC   initial and subsequent sets of combination oligonucleotides having a
CC   sequence region corresponding to the initial nucleotide sequence and a
CC   sequence region corresponding to the second oligonucleotide sequence.
CC   Also described is a method of creating a collection of recombination
CC   products between two genes. The methods and compositions of the present
CC   invention are useful in the field of synthetic gene technology, and more
CC   specifically, to generating a collection of recombination products
CC   between distinct nucleotide sequences. They can also be used in
CC   assembling a library, or a population or a collection of polypeptide
CC   variants that correspond to single or multiple polynucleotide
CC   recombination products. The present sequence represents an
CC   oligonucleotide used in the assembly of a synthetic antibody, which is
CC   used in an example from the present invention.
XX
SQ   Sequence 50 BP; 12 A; 11 C; 18 G; 9 T; 0 U; 0 Other;

  Query Match           14.5%;  Score 19.2;  DB 9;  Length 50;
  Best Local Similarity  75.0%;  Pred. No. 3.2e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         50 CCTGCATGCACACGTTCTGCGCGGCTTGCTAC 81
              |||| | | ||  |||| |||| |||| ||||
Db         44 CCTGAACGTACCAGTTCCGCGCCGCTTTCTAC 13


RESULT 3
AAZ68382
ID   AAZ68382 standard; DNA; 47 BP.
XX
AC   AAZ68382;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:2729.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation; diagnosis;
KW   single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB000822.
XX
PR   21-APR-1998;   98US-0082614P.
PR   23-NOV-1998;   98US-0109732P.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome.
XX
PS   Claim 3; Page 812; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the invention
CC   have a variety of uses: they can be used for high density mapping of the
CC   human genome, and in complex association studies and haplotyping studies
CC   which are useful in determining the genetic basis for disease states.
CC   Compositions and methods of the invention can also be useful for the
CC   identification of the targets for the development of pharmaceutical
CC   agents and diagnostic methods, as well as the characterisation of the
CC   differential efficacious responses to and side effects from
CC   pharmaceutical agents acting on a disease as well as other treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC   3367, are not actually given a sequence in the Sequence Listing from the
CC   present invention
XX
SQ   Sequence 47 BP; 11 A; 13 C; 10 G; 13 T; 0 U; 0 Other;

  Query Match           14.4%;  Score 19;  DB 3;  Length 47;
  Best Local Similarity  65.1%;  Pred. No. 3.7e+04;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         17 AGGACCTGCTGCACGACTGCGTGAGTTTGCAGCCCTGCATGCA 59
              ||| ||||| |||  |  |    |||||| |  ||| ||| ||
Db          5 AGGTCCTGCAGCAGTATGGTACCAGTTTGAAAGCCTCCATCCA 47


RESULT 4
ABN83043
ID   ABN83043 standard; DNA; 41 BP.
```

XX
AC ABN83043;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human ankyrin 12 probe 2.
XX
KW Human; ankyrin 12; cancer; HIV; probe; ss.
XX
OS Homo sapiens.
XX
PN CN1331128-A.
XX
PD 16-JAN-2002.
XX
PF 26-JUN-2000; 2000CN-00116733.
XX
PR 26-JUN-2000; 2000CN-00116733.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-305470/35.
XX
PT Polypeptide-human ankyrin 12 and polynucleotide for coding it.
XX
PS Example 6; Page 18; 31pp; Chinese.
XX
CC The sequence represents a probe used in the invention. The invention
CC relates to the novel human ankyrin 12 polypeptide. The polypeptide is
CC useful for treating diseases such as cancer and HIV infection. The
CC antagonist of the polypeptide and its medical action, and the application
CC of the polynucleotide are also disclosed
XX
SQ Sequence 41 BP; 8 A; 12 C; 11 G; 10 T; 0 U; 0 Other;

Query Match 14.2%; Score 18.8; DB 6; Length 41;
Best Local Similarity 76.7%; Pred. No. 4.1e+04;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 33 CTGCGTGAGTTTGCAGCCCTGCATGCACAC 62
||| |||||| ||| ||| | |||||||
Db 5 CTGTGTGAGTGAGCACCCCCGTGTGCACAC 34

RESULT 5
ADE03377/c
ID ADE03377 standard; DNA; 36 BP.
XX
AC ADE03377;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human immunoglobulin light chain PCR primer #57.
XX
KW antibody; platelet aggregation inhibition; platelet integrin receptor;
KW GPIIb/IIIa; activated thrombocyte; thrombosis; myocardial infarction;
KW primer; ss; human; PCR.
XX
OS Homo sapiens.
XX
PN EP1300419-A1.
XX
PD 09-APR-2003.
XX
PF 05-OCT-2001; 2001EP-00123851.
XX
PR 05-OCT-2001; 2001EP-00123851.
XX
PA (AFFI-) AFFIMED THERAPEUTICS AG.
XX
PI Buettner C, Schwarz M, Knackmuss S, Peter K, Roettgen P;
PI Little M;
XX
DR WPI; 2003-405595/39.
XX
PT New antibody, useful for preparing a composition for determining the
PT number of activated thrombocytes or for blocking the platelet integrin
PT receptor on thrombocytes for treating e.g., thrombosis or myocardial
PT infarction.
XX
PS Example 1; SEQ ID NO 107; 80pp; English.
XX
CC The invention comprises a human antibody for inhibiting platelet
CC aggregation by its exclusive binding to the activated state of platelet
CC integrin receptor GPIIb/IIIa. The antibody of the invention is useful for
CC preparing a diagnostic composition for determining the number of
CC activated thrombocytes or for blocking the platelet integrin receptor on
CC thrombocytes. The antibody of the invention is useful for treating
CC thrombosis or myocardial infarction. The present DNA sequence represents
CC a PCR primer that was used in an example of the invention.
XX
SQ Sequence 36 BP; 10 A; 12 C; 10 G; 4 T; 0 U; 0 Other;

Query Match 13.8%; Score 18.2; DB 9; Length 36;
Best Local Similarity 87.0%; Pred. No. 5.9e+04;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 46 CAGCCCTGCATGCACACGTTCTG 68
|||||||||| || || |||||||
Db 27 CAGCCCTGCCTGTACGCGTTCTG 5

RESULT 6
AAV60897
ID AAV60897 standard; DNA; 48 BP.
XX
AC AAV60897;
XX
DT 25-JAN-1999 (first entry)
XX
DE Oligonucleotide 910388 for constructing MAb Vh region variants.
XX
KW Mutation; mutagenesis; antigen-binding region; monoclonal antibody;
KW catalytic site; serine protease; complementarity determining region;
KW walk-through; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US5798208-A.
XX
PD 25-AUG-1998.
XX
PF 02-NOV-1992; 92US-00930600.
XX
PR 05-APR-1990; 90US-00505314.
XX
PA (CREA/) CREA R.
XX
PI Crea R;
XX
DR WPI; 1998-480376/41.
XX
PT Mutagenesis of pre-determined gene sequences - useful for systematic
PT changes of pre-determined amino acids to see their effect on protein
PT activity, and to create gene expression libraries.
XX
PS Example 1; Col 23; 33pp; English.
XX
CC The invention relates to a method of generating mutations in proteins by
CC synthesising a mixture of oligonucleotides in order to alter the codons
CC for specific amino acids within a defined region of the protein. Using a
CC range of oligonucleotides for the mutations, expression libraries of the
CC mutant protein can be constructed. As an example of the method, the

```
CC   antigen-binding region of the monoclonal antibody (MAb) MCPC603 (which
CC   binds phosphocholine) is altered to contain the catalytic triad residues
CC   for a serine protease. Specifically the amino acids to be altered are
CC   selected from the Asp of the complementarity determining region (CDR) 1
CC   region of the variable heavy chain (Vh) of the antibody, the His of Vh
CC   CDR3 and the Ser of the CDR2 from the light chain variable region (Vl).
CC   The mutagenesis is by a "walk-through" method. Oligonucleotides AAV60881-
CC   V60900 are used to assemble the Vh gene for monoclonal antibody MCPC603
CC   for the construction of variants (see AAV60860-V60880)
XX
SQ   Sequence 48 BP; 4 A; 11 C; 9 G; 12 T; 0 U; 12 Other;

  Query Match             13.8%;  Score 18.2;  DB 2;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 6.3e+04;
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    |  ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 7
ADD54968
ID   ADD54968 standard; DNA; 48 BP.
XX
AC   ADD54968;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Heavy chain variable region variant oligonucleotide #17.
XX
KW   ss; mutagenesis; protein mutagenesis; screening.
XX
OS   Synthetic.
XX
PN   US2003194807-A1.
XX
PD   16-OCT-2003.
XX
PF   20-FEB-2003; 2003US-00371404.
XX
PR   02-NOV-1992;   92US-00930600.
PR   30-MAY-1995;   95US-00453623.
XX
PA   (CREA/) CREA R.
XX
PI   Crea R;
XX
DR   WPI; 2003-844460/78.
XX
PT   Mutagenesis of a protein comprises introducing a predetermined amino acid
PT   into each set of selected sequence positions in a predefined region of
PT   the protein to produce a protein library comprising mutant proteins.
XX
PS   Example 1; Page 12; 44pp; English.
XX
CC   The invention relates to a method of mutagenesis of a protein. The
CC   methods are useful for generating libraries of mutant proteins that are
CC   of a practical size for screening, for studying the role of amino acids
CC   in protein structure and function and for developing new or improved
CC   proteins and polypeptides such as enzymes, antibodies their binding
CC   fragments or analogues. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 48 BP; 4 A; 11 C; 9 G; 12 T; 0 U; 12 Other;

  Query Match             13.8%;  Score 18.2;  DB 9;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 6.3e+04;
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    |  ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 8
AAQ42415/c
ID   AAQ42415 standard; cDNA; 47 BP.
XX
AC   AAQ42415;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1993  (first entry)
XX
DE   IDUA primer ID47.
XX
KW   Alpha-L-iduronidase; glycosaminoglycan alpha-L-iduronohydrolase;
KW   mucopolysaccharidosis type I; MPS-1; primer; probe; ss.
XX
OS   Synthetic.
XX
PN   WO9310244-A1.
XX
PD   27-MAY-1993.
XX
PF   12-NOV-1992;   92WO-AU000611.
XX
PR   14-NOV-1991;   91AU-00009490.
XX
PA   (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX
PI   Scott HS,  Anson DS,  Orsborn AM,  Nelson PV,  Clements PR;
PI   Morris CP,  Hopwood JJ;
XX
DR   WPI; 1993-182558/22.
XX
PT   Isolated nucleic acid molecule for alpha-L-idoronidase hydrolase
PT   deficiency diagnosis - comprises sequence of nucleotide(s) encoding or
PT   complementary to mammalian alpha-L-idoronidase hydrolase fragment, for
PT   treating lysosomal storage disorder.
XX
PS   Example; Page 14; 47pp; English.
XX
CC   The IDUA nucleic acid was isolated using oligonucleotides synthesised on
CC   the basis of IDUA amino acid sequences. Primer ID47 was based on residues
CC   2 to 7 of peptide 8 (see AAR37425). Primer ID13 was based on residues 13
CC   to 7 of the 74/13 kD amino terminal (see AAR37425). (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 47 BP; 14 A; 16 C; 11 G; 6 T; 0 U; 0 Other;

  Query Match             13.6%;  Score 18;  DB 2;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 7.1e+04;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
              |||||  |    | | ||||  ||||||| | |  ||  ||||
Db         47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 9
ABS56353/c
ID   ABS56353 standard; DNA; 48 BP.
XX
AC   ABS56353;
XX
DT   20-JAN-2003  (first entry)
XX
DE   PCR primer, TPOVHP6, used to construct an MPL binding sequence of TPO.
XX
KW   Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
KW   complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
KW   thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
KW   T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
KW   proliferation; growth; differentiation; haematopoietic cell; antibody;
KW   platelet progenitor cell; immune disorder; thrombocytopenia; primer;
```

KW disseminated intravascular coagulation; stem cell; transplantation;
KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant;
KW consensus variable heavy chain domain; CONVH; PCR; ss; PCR knitting.
XX
OS Synthetic.
XX
PN WO200278612-A2.
XX
PD 10-OCT-2002.
XX
PF 02-APR-2002; 2002WO-US010301.
XX
PR 02-APR-2001; 2001US-0281183P.
XX
PA (PURD) PURDUE PHARMA LP.
XX
PI Soltis DA, Burch RM, Ogert RA;
XX
DR WPI; 2003-040615/03.
XX
PT New thrombopoietin synthebodies, useful for stimulating proliferation,
PT growth, or differentiation of hematopoietic cells, for treating or
PT preventing hematopoietic or immune disorders, e.g. thrombocytopenia.
XX
PS Example 1; Page 72; 97pp; English.
XX
CC The invention discloses a variant of an immunoglobulin (Ig) variable
CC heavy or light chain domain that comprises at least one complementarity
CC determining region (CDR) and framework regions flanking the CDR. The CDR
CC also has added or substituted to it, at least one binding sequence which
CC is heterologous to the CDR and is an antigenic, agonistic sequence from a
CC thrombopoietin (TPO) receptor (MPL) binding sequence. The antigenic
CC sequence can be a binding sequence heterologous to the CDR, a cytotoxic T
CC -lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
CC cell sequence or a combination of each. The variant or thrombopoietin
CC synthebody, pharmaceutical and vaccine compositions are useful for
CC stimulating proliferation, growth or differentiation of haematopoietic
CC cells, particularly platelet progenitor cells. The variants are also
CC useful for treating or preventing haematopoietic or immune disorders
CC resulting from chemotherapy, radiation therapy, or bone marrow
CC transfusions (e.g. thrombocytopenia or disseminated intravascular
CC coagulation). Compositions comprising the synthebodies can be used for
CC the mobilisation, amplification and ex vivo expansion of stem cells and
CC committed precursor cells for autologous and allogeneic transplantation
CC as well as for the expansion of stem cells for gene therapy. They are
CC also useful as diagnostic or analytical reagents for studying the
CC function of thrombopoietin and its receptor in vivo or in vitro. The
CC sequence presented is the PCR primer, TPOVHP6, which was used to
CC construct a variable heavy chain region gene containing the MPL binding
CC sequence of TPO using the PCR knitting technique
XX
SQ Sequence 48 BP; 10 A; 14 C; 15 G; 9 T; 0 U; 0 Other;

Query Match 13.6%; Score 18; DB 7; Length 48;
Best Local Similarity 64.3%; Pred. No. 7.2e+04;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 71 CGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTC 112
|| | || ||||| | |||| ||| || | ||||| | |
Db 48 CTGTTTACTACTGCGCTAGGATCGAGGGCCCCACCCTGAGAC 7

RESULT 10
ABZ01079
ID ABZ01079 standard; DNA; 50 BP.
XX
AC ABZ01079;
XX
DT 09-JAN-2003 (first entry)
XX
DE Human leukocyte gene expression profiling probe SEQ ID NO 1070.
XX
KW T7; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200257414-A2.
XX
PD 25-JUL-2002.
XX
PF 22-OCT-2001; 2001WO-US047856.
XX
PR 20-OCT-2000; 2000US-0241994P.
PR 08-JUN-2001; 2001US-0296764P.
XX
PA (BIOC-) BIOCARDIA INC.
XX
PI Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quertermous T, Johnson F;
XX
DR WPI; 2002-636525/68.
XX
PT New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS Claim 1; Page 359; 0pp; English.
XX
CC The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC predicting disease complications in an individual or monitoring response
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ Sequence 50 BP; 11 A; 12 C; 9 G; 18 T; 0 U; 0 Other;

Query Match 13.6%; Score 18; DB 6; Length 50;
Best Local Similarity 70.6%; Pred. No. 7.2e+04;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 51 CTGCATGCACACGTTCTGCGCGGCTTGCTACTCG 84
||| || || |||||| | |||||||| | |
Db 10 CTGTGTGAACTGTTTCTGCACTGCTTGCTAATGG 43

RESULT 11
AAL31778/c
ID AAL31778 standard; DNA; 50 BP.
XX
AC AAL31778;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #4986.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX

PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Claim 1; Page 2821; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-
CC protein coupled receptors and thioesterases. The present sequence is one
CC such oligonucleotide. The oligonucleotides and the peptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of the proteins listed above.
CC Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms
XX
SQ Sequence 50 BP; 15 A; 13 C; 14 G; 8 T; 0 U; 0 Other;

Query Match 13.5%; Score 17.8; DB 4; Length 50;
Best Local Similarity 62.2%; Pred. No. 8.2e+04;
Matches 28; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 63 GTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCT 107
Db 48 GCTCAGAGCCTCTAGATTCACAGTCTGCATGGTGGGCTTCTGCCT 4

RESULT 12
AAV02057
ID AAV02057 standard; cDNA; 42 BP.
XX
AC AAV02057;
XX
DT 26-JUN-1998 (first entry)
XX
DE Mutated hCS peptide coding sequence.
XX
KW Corticostatin; somatostatin; hCS-17 peptide; hormone-producing tumour;
KW gastric ulcer; dementia; growth disorder; hormone secretion regulation;
KW digestive system regulation; neural inhibitor; therapy; ds.
XX
OS Synthetic.
XX
PN WO9746668-A1.
XX
PD 11-DEC-1997.
XX
PF 05-JUN-1997; 97WO-JP001911.
XX
PR 07-JUN-1996; 96JP-00146052.
PR 19-SEP-1996; 96JP-00247710.
PR 15-OCT-1996; 96JP-00272422.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Fukusumi S, Kitada C;
XX
DR WPI; 1998-042177/04.
DR P-PSDB; AAW44022.
XX
PT Peptide having corticostatin or somatostatin or activity - useful as
PT anticancer and antiulcer agent, and for control of dementia and growth
PT abnormalities.
XX
PS Claim 14; Page 138; 174pp; Japanese.
XX
CC This sequence encodes a mutated hCS peptide, which is a peptide of the
CC invention, and has corticostatin or somatostatin activity. Antibodies
CC recognising hCS-17 can be used to screen for a compound that modulates,
CC i.e. an agonist or antagonist, the binding of hCS-17 to its receptor, and
CC to assay for hCS-17, e.g. diagnosis. hCS-17, the DNA encoding it or a
CC receptor agonist or antagonist can be used to treat and prevent hormone-
CC producing tumours (e.g. tumours producing gastrin or insulin), gastric
CC ulcers and dementia, regulate sleep and control growth disorders (e.g.
CC acromegaly, gigantism and dwarfism). They can also be used to regulate
CC hormone secretion and the digestive system (e.g. to treat diabetes), and
CC as tumour multiplication or neural inhibitors
XX
SQ Sequence 42 BP; 7 A; 15 C; 8 G; 12 T; 0 U; 0 Other;

Query Match 13.3%; Score 17.6; DB 2; Length 42;
Best Local Similarity 65.0%; Pred. No. 9.1e+04;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 48 GCCCTGCATGCACACGTTCTGCGCGGCTTGCTACTCGGGC 87
Db 3 GCCCTGCAGGAACTTCTTCTGGAAGACCTTCTCCTCCTGC 42

RESULT 13
AAV02046
ID AAV02046 standard; cDNA; 45 BP.
XX
AC AAV02046;
XX
DT 26-JUN-1998 (first entry)
XX
DE hCS-15 peptide coding sequence.
XX
KW Corticostatin; somatostatin; hCS-15 peptide; hormone-producing tumour;
KW gastric ulcer; dementia; growth disorder; hormone secretion regulation;
KW digestive system regulation; neural inhibitor; therapy; ds.
XX
OS Synthetic.
XX
PN WO9746668-A1.
XX
PD 11-DEC-1997.
XX
PF 05-JUN-1997; 97WO-JP001911.
XX
PR 07-JUN-1996; 96JP-00146052.
PR 19-SEP-1996; 96JP-00247710.
PR 15-OCT-1996; 96JP-00272422.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Fukusumi S, Kitada C;
XX
DR WPI; 1998-042177/04.
DR P-PSDB; AAW44019.
XX
PT Peptide having corticostatin or somatostatin or activity - useful as
PT anticancer and antiulcer agent, and for control of dementia and growth

PT abnormalities.
XX
PS Claim 11; Page 119; 174pp; Japanese.
XX
CC This sequence encodes the peptide hCS-15. hCS-15 is a peptide of the
CC invention, and has corticostatin or somatostatin activity. Antibodies
CC recognising hCS-17 can be used to screen for a compound that modulates,
CC i.e. an agonist or antagonist, the binding of hCS-17 to its receptor, and
CC to assay for hCS-17, e.g. diagnosis. hCS-17, the DNA encoding it or a
CC receptor agonist or antagonist can be used to treat and prevent hormone-
CC producing tumours (e.g. tumours producing gastrin or insulin), gastric
CC ulcers and dementia, regulate sleep and control growth disorders (e.g.
CC acromegaly, gigantism and dwarfism). They can also be used to regulate
CC hormone secretion and the digestive system (e.g. to treat diabetes), and
CC as tumour multiplication or neural inhibitors
XX
SQ Sequence 45 BP; 10 A; 15 C; 8 G; 12 T; 0 U; 0 Other;

Query Match 13.3%; Score 17.6; DB 2; Length 45;
Best Local Similarity 65.0%; Pred. No. 9.2e+04;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 48 GCCCTGCATGCACACGTTCTGCGCGGCTTGCTACTCGGGC 87
||||||| | || ||||| | | | || ||| ||
Db 3 GCCCTGCAGGAACTTCTTCTGGAAGACCTTCTCCTCCTGC 42

RESULT 14
AAV02056
ID AAV02056 standard; cDNA; 48 BP.
XX
AC AAV02056;
XX
DT 26-JUN-1998 (first entry)
XX
DE Mutated hCS peptide coding sequence.
XX
KW Corticostatin; somatostatin; hCS-17 peptide; hormone-producing tumour;
KW gastric ulcer; dementia; growth disorder; hormone secretion regulation;
KW digestive system regulation; neural inhibitor; therapy; ds.
XX
OS Synthetic.
XX
PN WO9746668-A1.
XX
PD 11-DEC-1997.
XX
PF 05-JUN-1997; 97WO-JP001911.
XX
PR 07-JUN-1996; 96JP-00146052.
PR 19-SEP-1996; 96JP-00247710.
PR 15-OCT-1996; 96JP-00272422.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Fukusumi S, Kitada C;
XX
DR WPI; 1998-042177/04.
DR P-PSDB; AAW44021.
XX
PT Peptide having corticostatin or somatostatin or activity - useful as
PT anticancer and antiulcer agent, and for control of dementia and growth
PT abnormalities.
XX
PS Claim 14; Page 138; 174pp; Japanese.
XX
CC This sequence encodes a mutated hCS peptide, which is a peptide of the
CC invention, and has corticostatin or somatostatin activity. Antibodies
CC recognising hCS-17 can be used to screen for a compound that modulates,
CC i.e. an agonist or antagonist, the binding of hCS-17 to its receptor, and
CC to assay for hCS-17, e.g. diagnosis. hCS-17, the DNA encoding it or a
CC receptor agonist or antagonist can be used to treat and prevent hormone-
CC producing tumours (e.g. tumours producing gastrin or insulin), gastric
CC ulcers and dementia, regulate sleep and control growth disorders (e.g.
CC acromegaly, gigantism and dwarfism). They can also be used to regulate
CC hormone secretion and the digestive system (e.g. to treat diabetes), and
CC as tumour multiplication or neural inhibitors
XX
SQ Sequence 48 BP; 10 A; 16 C; 10 G; 12 T; 0 U; 0 Other;

Query Match 13.3%; Score 17.6; DB 2; Length 48;
Best Local Similarity 65.0%; Pred. No. 9.3e+04;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 48 GCCCTGCATGCACACGTTCTGCGCGGCTTGCTACTCGGGC 87
||||||| | || ||||| | | | || ||| ||
Db 9 GCCCTGCAGGAACTTCTTCTGGAAGACCTTCTCCTCCTGC 48

RESULT 15
AAL34532
ID AAL34532 standard; DNA; 50 BP.
XX
AC AAL34532;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #7740.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Claim 1; Page 3626; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-
CC protein coupled receptors and thioesterases. The present sequence is one
CC such oligonucleotide. The oligonucleotides and the peptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of the proteins listed above.
CC Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic

CC organisms
XX
SQ Sequence 50 BP; 12 A; 9 C; 14 G; 15 T; 0 U; 0 Other;

Query Match 13.3%; Score 17.6; DB 4; Length 50;
Best Local Similarity 65.0%; Pred. No. 9.4e+04;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy         30 CGACTGCGTGAGTTTGCAGCCCTGCATGCACACGTTCTGC 69
              | |||| |||||| ||     ||||   |||  ||| |||
Db          6 CAACTGTGTGAGTGTGATAAGCTGCTGCCACCTGTTTTGC 45
```

Search completed: May 26, 2004, 05:53:46
Job time : 114.477 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:16:49 ; Search time 25.4982 Seconds
(without alignments)
2872.886 Million cell updates/sec

Title: US-10-048-046-1_COPY_997_1128
Perfect score: 132
Sequence: 1 acatgcatcatctgccagga.....ctacctgccgctgtcccgtg 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20.2	15.3	47	4	US-09-422-978-2729	Sequence 2729, Ap
c	2	19.2	14.5	47	1	US-08-236-311-18	Sequence 18, Appl
c	3	19.2	14.5	47	3	US-08-457-918-18	Sequence 18, Appl
	4	19	14.4	47	1	US-08-236-311-15	Sequence 15, Appl
	5	19	14.4	47	3	US-08-457-918-15	Sequence 15, Appl
	6	18.2	13.8	48	2	US-08-452-724A-20	Sequence 20, Appl
	7	18.2	13.8	48	4	US-08-453-623-20	Sequence 20, Appl
c	8	18	13.6	45	2	US-08-602-716A-8	Sequence 8, Appli
c	9	18	13.6	47	3	US-09-191-171-1	Sequence 1, Appli
c	10	18	13.6	47	3	US-09-385-707-1	Sequence 1, Appli
c	11	18	13.6	47	4	US-09-639-696C-1	Sequence 1, Appli
c	12	17.4	13.2	32	4	US-09-581-105-9	Sequence 9, Appli
	13	17.4	13.2	40	6	5240847-33	Patent No. 5240847
c	14	17.2	13.0	46	5	PCT-US95-13975-35	Sequence 35, Appl
c	15	17.2	13.0	46	6	5196308-1	Patent No. 5196308
	16	17	12.9	36	1	US-08-425-299A-5	Sequence 5, Appli
	17	17	12.9	36	3	US-08-895-601-15	Sequence 15, Appl
	18	17	12.9	36	3	US-08-767-942A-28	Sequence 28, Appl
c	19	17	12.9	47	4	US-09-422-978-735	Sequence 735, App
	20	17	12.9	48	1	US-08-171-389-228	Sequence 228, App
	21	17	12.9	48	1	US-08-391-000-11	Sequence 11, Appl
	22	17	12.9	48	1	US-08-123-936-228	Sequence 228, App
	23	17	12.9	48	2	US-08-741-931-11	Sequence 11, Appl
	24	17	12.9	48	2	US-08-475-228A-228	Sequence 228, App
	25	17	12.9	48	3	US-08-482-080A-228	Sequence 228, App
	26	17	12.9	48	4	US-09-354-947-228	Sequence 228, App
	27	17	12.9	48	5	PCT-US93-12388-228	Sequence 228, App
c	28	17	12.9	50	3	US-08-959-212-8	Sequence 8, Appli
c	29	16.8	12.7	30	1	US-08-426-819A-27	Sequence 27, Appl
	30	16.8	12.7	40	4	US-09-262-537-51	Sequence 51, Appl
c	31	16.8	12.7	45	1	US-07-601-094-6	Sequence 6, Appli
c	32	16.8	12.7	45	1	US-08-012-735-6	Sequence 6, Appli
c	33	16.8	12.7	47	1	US-08-285-440-17	Sequence 17, Appl
c	34	16.8	12.7	47	1	US-08-630-349-17	Sequence 17, Appl
c	35	16.8	12.7	50	1	US-08-171-389-551	Sequence 551, App
c	36	16.8	12.7	50	1	US-08-123-936-551	Sequence 551, App
c	37	16.8	12.7	50	2	US-08-475-228A-551	Sequence 551, App
c	38	16.8	12.7	50	3	US-08-482-080A-551	Sequence 551, App
c	39	16.8	12.7	50	4	US-09-354-947-551	Sequence 551, App
c	40	16.8	12.7	50	5	PCT-US93-12388-551	Sequence 551, App
c	41	16.6	12.6	42	4	US-09-442-021-24	Sequence 24, Appl
	42	16.6	12.6	47	4	US-09-422-978-2420	Sequence 2420, Ap
c	43	16.6	12.6	50	1	US-08-171-389-400	Sequence 400, App
c	44	16.6	12.6	50	1	US-08-123-936-400	Sequence 400, App
c	45	16.6	12.6	50	2	US-08-475-228A-400	Sequence 400, App

ALIGNMENTS

RESULT 1
US-09-422-978-2729
; Sequence 2729, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CP1
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2729
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-16284-389 : polymorphic base T or C
US-09-422-978-2729

Query Match 15.3%; Score 20.2; DB 4; Length 47;
Best Local Similarity 65.1%; Pred. No. 2.3e+03;
Matches 28; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

```
Qy         17 AGGACCTGCTGCACGACTGCGTGAGTTTGCAGCCCTGCATGCA 59
              ||| ||||| ||| | |:  ||||||| |  ||| ||| ||
Db          5 AGGTCCTGCAGCAGTATGGYACCAGTTTGAAAGCCTCCATCCA 47
```

RESULT 2
US-08-236-311-18/c
; Sequence 18, Application US/08236311
; Patent No. 5565335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adheson Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd

```
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/236,311
;       FILING DATE:  02-MAY-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/936190
;       FILING DATE:  26-AUG-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/842777
;       FILING DATE:  18-FEB-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/250785
;       FILING DATE:  28-SEP-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/104329
;       FILING DATE:  02-OCT-1987
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hasak, Janet E.
;       REGISTRATION NUMBER:  28,616
;       REFERENCE/DOCKET NUMBER:  444P1C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-1896
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-236-311-18

  Query Match             14.5%;  Score 19.2;  DB 1;  Length 47;
  Best Local Similarity   67.5%;  Pred. No. 4.5e+03;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCC 113
              |||||| || | |||| |    |||||| |||| |   ||
Db         46 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCC 7


RESULT 3
US-08-457-918-18/c
; Sequence 18, Application US/08457918
; Patent No. 6117655
;  GENERAL INFORMATION:
;    APPLICANT:  Capon, Daniel J.
;    APPLICANT:  Gregory, Timothy J.
;    TITLE OF INVENTION:  Adheson Variants
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/457,918
;      FILING DATE:  1-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/236311
;      FILING DATE:  02-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/936190
;      FILING DATE:  26-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/842777
;      FILING DATE:  18-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/250785
;      FILING DATE:  28-SEP-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/104329
;      FILING DATE:  02-OCT-1987
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kubinec, Jeffrey S.
;      REGISTRATION NUMBER:  36,575
;      REFERENCE/DOCKET NUMBER:  P0444P1C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8228
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-457-918-18

  Query Match             14.5%;  Score 19.2;  DB 3;  Length 47;
  Best Local Similarity   67.5%;  Pred. No. 4.5e+03;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCC 113
              |||||| || | |||| |    |||||| |||| |   ||
Db         46 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCC 7


RESULT 4
US-08-236-311-15
; Sequence 15, Application US/08236311
; Patent No. 5565335
;  GENERAL INFORMATION:
;    APPLICANT:  Capon, Daniel J.
;    APPLICANT:  Gregory, Timothy J.
;    TITLE OF INVENTION:  Adheson Variants
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/236,311
;      FILING DATE:  02-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/936190
;      FILING DATE:  26-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/842777
;      FILING DATE:  18-FEB-1992
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/250785
;      FILING DATE:  28-SEP-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/104329
;      FILING DATE:  02-OCT-1987
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hasak, Janet E.
;      REGISTRATION NUMBER:  28,616
;      REFERENCE/DOCKET NUMBER:  444P1C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1896
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-236-311-15

  Query Match             14.4%;  Score 19;  DB 1;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 5.2e+03;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTG 108
              |||||| || | |||| |    |||||| |||| |
Db         12 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAG 46


RESULT 5
US-08-457-918-15
; Sequence 15, Application US/08457918
; Patent No. 6117655
;  GENERAL INFORMATION:
;    APPLICANT:  Capon, Daniel J.
;    APPLICANT:  Gregory, Timothy J.
;    TITLE OF INVENTION:  Adheson Variants
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/457,918
;      FILING DATE:  1-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/236311
;      FILING DATE:  02-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/936190
;      FILING DATE:  26-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/842777
;      FILING DATE:  18-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/250785
;      FILING DATE:  28-SEP-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/104329
;      FILING DATE:  02-OCT-1987
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kubinec, Jeffrey S.
;      REGISTRATION NUMBER:  36,575
;      REFERENCE/DOCKET NUMBER:  P0444P1C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8228
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-457-918-15

  Query Match             14.4%;  Score 19;  DB 3;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 5.2e+03;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTG 108
              |||||| || | |||| |    |||||| |||| |
Db         12 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAG 46


RESULT 6
US-08-452-724A-20
; Sequence 20, Application US/08452724A
; Patent No. 5830650
;  GENERAL INFORMATION:
;    APPLICANT:  Crea, Roberto
;    TITLE OF INVENTION:  Walk-Through Mutagenesis
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/452,724A
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,600
;      FILING DATE:  05-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/02362
;      FILING DATE:  05-APR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/505,314
;      FILING DATE:  05-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brook Esq., David E.
;      REGISTRATION NUMBER:  22,592
;      REFERENCE/DOCKET NUMBER:  RC90-01AZ
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
US-08-452-724A-20

  Query Match             13.8%;  Score 18.2;  DB 2;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 8.8e+03;
```

```
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    | ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 7
US-08-453-623-20
; Sequence 20, Application US/08453623
; Patent No. 6649340
;  GENERAL INFORMATION:
;       APPLICANT: Crea, Roberto
;       TITLE OF INVENTION: Walk-Through Mutagenesis
;       NUMBER OF SEQUENCES: 59
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;            STREET: 2 Militia Drive
;            CITY: Lexington
;            STATE: MA
;            COUNTRY: USA
;            ZIP: 02173
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/453,623
;            FILING DATE: 30-May-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/930,600
;            FILING DATE: 05-APR-1991
;            APPLICATION NUMBER: PCT/US91/02362
;            FILING DATE: 05-APR-1991
;            APPLICATION NUMBER: US 07/505,314
;            FILING DATE: 05-APR-1990
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Brook, David E.
;            REGISTRATION NUMBER: 22,592
;            REFERENCE/DOCKET NUMBER: RC90-01AY
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 861-6240
;            TELEFAX: (617) 861-9540
;   INFORMATION FOR SEQ ID NO: 20:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 48 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: unknown
;       SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-453-623-20

  Query Match             13.8%;  Score 18.2;  DB 4;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 8.8e+03;
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    | ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 8
US-08-602-716A-8/c
; Sequence 8, Application US/08602716A
; Patent No. 5962664
;  GENERAL INFORMATION:
;    APPLICANT:  FRIEDHOFF, Arnold J.
;    APPLICANT:  BASHAM, Daryl A.
;    APPLICANT:  MILLER, Jeanette C.
;    TITLE OF INVENTION:  PSYCHOSIS PROTECTING NUCLEIC ACID,
;    TITLE OF INVENTION:  PEPTIDES, COMPOSITIONS AND METHODS OF USE
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/602,716A
;      FILING DATE:  23-FEB-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/060,560
;      FILING DATE:  13-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05545
;      FILING DATE:  13-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  FRIEDHOFF=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-602-716A-8

  Query Match             13.6%;  Score 18;  DB 2;  Length 45;
  Best Local Similarity   80.8%;  Pred. No. 9.9e+03;
  Matches   21;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGC 99
              |||| | | | ||||||||||||| ||
Db         29 CTTGTTGCCCAGGCTGGATGGAGTGC 4


RESULT 9
US-09-191-171-1/c
; Sequence 1, Application US/09191171
; Patent No. 6149909
; Patent No. 6149909 6143294
;  GENERAL INFORMATION:
;    APPLICANT:  SCOTT, Hamish S.
;    APPLICANT:  ANSON, Donald S.
;    APPLICANT:  ORSBORN, Annette M.
;    APPLICANT:  NELSON, Paul V.
;    APPLICANT:  CLEMENTS, Peter R.
;    APPLICANT:  MORRIS, Charles P.
;    APPLICANT:  HOPWOOD, John J.
;    TITLE OF INVENTION:  SYNTHETIC ALPHA-L-IDURONIDASE AND GENETIC
;    TITLE OF INVENTION:  SEQUENCES ENCODING SAME
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY SCOTT MURPHY & PRESSER
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/191,171
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/494,104
;       FILING DATE:  23-JUN-1995
;       APPLICATION NUMBER:  AU PK9490/91
;       FILING DATE:  14-NOV-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  AU PCT/AU92/00611
;       FILING DATE:  12-NOV-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/084,254
;       FILING DATE:  07-JUL-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DiGiglio, Frank S.
;       REGISTRATION NUMBER:  31,346
;       REFERENCE/DOCKET NUMBER:  8978Z
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  516-742-4343
;       TELEFAX:  516-742-4366
;       TELEX:  230 901 SANS UR
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Oligonucleotide DNA
US-09-191-171-1

  Query Match             13.6%;  Score 18;  DB 3;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 1e+04;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
               ||||| |    | | ||||  ||||||| | |   ||  ||||
Db          47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 10
US-09-385-707-1/c
; Sequence 1, Application US/09385707
; Patent No. 6238662
;  GENERAL INFORMATION:
;     APPLICANT:  SCOTT, Hamish S.
;     APPLICANT:  ANSON, Donald S.
;     APPLICANT:  ORSBORN, Annette M.
;     APPLICANT:  NELSON, Paul V.
;     APPLICANT:  CLEMENTS, Peter R.
;     APPLICANT:  MORRIS, Charles P.
;     APPLICANT:  HOPWOOD, John J.
;     TITLE OF INVENTION:  SYNTHETIC ALPHA-L-IDURONIDASE AND GENETIC
;     TITLE OF INVENTION:  SEQUENCES ENCODING SAME
;     NUMBER OF SEQUENCES:  7
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  SCULLY SCOTT MURPHY & PRESSER
;       STREET:  400 Garden City Plaza
;       CITY:  Garden City
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  11530
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/385,707
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/494,104
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DiGiglio, Frank S.
;       REGISTRATION NUMBER:  31,346
;       REFERENCE/DOCKET NUMBER:  8978Z
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  516-742-4343
;       TELEFAX:  516-742-4366
;       TELEX:  230 901 SANS UR
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  47 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Oligonucleotide DNA
US-09-385-707-1

  Query Match             13.6%;  Score 18;  DB 3;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 1e+04;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
               ||||| |    | | ||||  ||||||| | |   ||  ||||
Db          47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 11
US-09-639-696C-1/c
; Sequence 1, Application US/09639696C
; Patent No. 6524835
;   GENERAL INFORMATION:
;       APPLICANT: SCOTT, Hamish S.
;                  ANSON, Donald S.
;                  ORSBORN, Annette M.
;                  NELSON, Paul V.
;                  CLEMENTS, Peter R.
;                  MORRIS, Charles P.
;                  HOPWOOD, John J.
;       TITLE OF INVENTION: SYNTHETIC ALPHA-L-IDURONIDASE AND GENETIC
;                           SEQUENCES ENCODING SAME
;       NUMBER OF SEQUENCES: 6
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON PEABODY LLP
;            STREET: 990 Stewart Avenue
;            CITY: Garden City
;            STATE: New York
;            COUNTRY: USA
;            ZIP: 11530
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/639,696C
;            FILING DATE: 16-Aug-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 09/191,171
;            FILING DATE: 13-NOV-1998
;            APPLICATION NUMBER: US 08/494,104
;            FILING DATE: 23-JUN-1995
;            APPLICATION NUMBER: AU PK9490/91
;            FILING DATE: 14-NOV-1991
;            APPLICATION NUMBER: AU PCT/AU92/00611
;            FILING DATE: 12-NOV-1992
;            APPLICATION NUMBER: US 08/084,254
;            FILING DATE: 07-JUL-1993
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Pokalsky, Ann R.
;            REGISTRATION NUMBER: 34,697
;            REFERENCE/DOCKET NUMBER: 2249/304
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 516-832-7572
;            TELEFAX: 516-832-7555
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 47 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: Oligonucleotide DNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-639-696C-1

  Query Match             13.6%;  Score 18;  DB 4;  Length 47;
  Best Local Similarity   64.3%;  Pred. No. 1e+04;
  Matches   27;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          61 ACGTTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
               ||||| |    | | ||||  ||||||| | |   ||  ||||
Db          47 ACGTTGTCGAAGTCGTGCTGGTCGGGCTCGTTCCAGGTCTCG 6


RESULT 12
US-09-581-105-9/c
; Sequence 9, Application US/09581105
; Patent No. 6603062
; GENERAL INFORMATION:
;  APPLICANT:  Schmidt, et al.
;  TITLE OF INVENTION:   Method For Producing Transgenic Plants With Modified 5-Aminolev
;  TITLE OF INVENTION:  Acid Biosynthesis, And Method For Identifying 5-Aminolevulinic A
;  TITLE OF INVENTION:  Effectors
;  FILE REFERENCE:   514413-3827
;  CURRENT APPLICATION NUMBER: US/09/581,105
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:   PCT/EP98/08028
;  PRIOR FILING DATE:   1998-12-10
;  PRIOR APPLICATION NUMBER:   GERMANY 197 54 929.2
;  PRIOR FILING DATE:   1997-12-10
;  NUMBER OF SEQ ID NOS:   10
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: primer ALAS1
US-09-581-105-9

  Query Match             13.2%;  Score 17.4;  DB 4;  Length 32;
  Best Local Similarity   77.8%;  Pred. No. 1.4e+04;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          35 GCGTGAGTTTGCAGCCCTGCATGCACA 61
               | |  || ||| || ||||||||||||
Db          30 GTGCCAGATTGTAGTCCTGCATGCACA 4


RESULT 13
5240847-33
;Patent No. 5240847
;    APPLICANT: HECKL, KONRAD;SPEVAK, WALTER;OSTERMANN, ELINBORG;
;ZOPHEL, ANDREAS;KRYSTEK, EDELTRAUD;MAURER-FOGY, INGRID;
;WICHE-CASTANON, MARIA J.;STRATOWA, CHRISTIAN;HAUPTMANN, RUDOLF
;    TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE
;(HMN-SOD)
;    NUMBER OF SEQUENCES: 34
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/167,261
;      FILING DATE: 11-MAR-1988
;SEQ ID NO:33:
;      LENGTH: 40
5240847-33

  Query Match             13.2%;  Score 17.4;  DB 6;  Length 40;
  Best Local Similarity   68.6%;  Pred. No. 1.4e+04;
  Matches   24;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 ATGCATCATCTGCCAGGACCTGCTGCACGACTGCG 37
              | |||   | ||||| ||| |||   |||||| ||
Db          1 AAGCACTCTTTGCCAAGACTTGCCATACGACTACG 35


RESULT 14
PCT-US95-13975-35/c
; Sequence 35, Application PC/TUS9513975
;  GENERAL INFORMATION:
;    APPLICANT:  Cochran, Mark D.
;    APPLICANT:  McDonell, Michael W.
;    TITLE OF INVENTION:  Recombinant Feline Herpes virus
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM 330 466 DX2
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13975
;      FILING DATE:  26-OCT-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/329,883
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  39118-PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)278-0400
;      TELEFAX:  (212)391-0525
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Human cytomegalovirus; Feline leukemia virus
PCT-US95-13975-35

  Query Match             13.0%;  Score 17.2;  DB 5;  Length 46;
  Best Local Similarity   65.8%;  Pred. No. 1.7e+04;
  Matches   25;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          82 TCGGGCTGGATGGAGCGCTCGTCCCTGTGTCCTACCTG 119
               ||||||   |||| |  | ||||| ||||| |   ||
Db          41 TCGGGCCCATTGGATCCATGGTCCCGGTGTCTTCTATG 4


RESULT 15
5196308-1/c
;Patent No. 5196308
```

; APPLICANT: NEPOM, GERALD T.;NEPOM, BARBARA S.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING THE DQW3.2
;ALLELE ASSOCIATED WITH INCREASED RISK OF INSULIN-DEPENDENT
;DIABETES MELLITUS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/325,058
; FILING DATE: 13-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 745,321
; FILING DATE: 14-JUN-1985
;SEQ ID NO:1:
; LENGTH: 46
5196308-1

Query Match 13.0%; Score 17.2; DB 6; Length 46;
Best Local Similarity 73.3%; Pred. No. 1.7e+04;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy         87 CTGGATGGAGCGCTCGTCCCTGTGTCCTAC 116
              |||||||| | | | ||| ||||| || |
Db         30 CTGGATGGGGAGATGGTCACTGTGGGCTCC 1
```

Search completed: May 26, 2004, 09:20:41
Job time : 26.4982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:41:55 ; Search time 576.198 Seconds
(without alignments)
1041.142 Million cell updates/sec

Title: US-10-048-046-1_COPY_997_1128
Perfect score: 132
Sequence: 1 acatgcatcatctgccagga.....ctacctgccgctgtcccgtg 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 1612472

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20.2	15.3	47	16	US-10-349-143-2729	Sequence 2729, Ap
c	2	19.2	14.5	47	15	US-10-097-044A-18	Sequence 18, Appl
	3	19	14.4	47	15	US-10-097-044A-15	Sequence 15, Appl
	4	18.2	13.8	48	15	US-10-371-404-20	Sequence 20, Appl
	5	18	13.6	50	16	US-10-131-827-1070	Sequence 1070, Ap
c	6	17.8	13.5	44	15	US-10-076-802-81	Sequence 81, Appl
c	7	17.8	13.5	44	15	US-10-076-802-110	Sequence 110, App
c	8	17.8	13.5	44	15	US-10-076-802-111	Sequence 111, App
c	9	17.8	13.5	44	16	US-10-367-169-81	Sequence 81, Appl
c	10	17.8	13.5	44	16	US-10-367-169-110	Sequence 110, App
c	11	17.8	13.5	44	16	US-10-367-169-111	Sequence 111, App
c	12	17.4	13.2	32	13	US-10-458-568-9	Sequence 9, Appli
	13	17.4	13.2	49	15	US-10-216-054A-14	Sequence 14, Appl
	14	17.4	13.2	49	15	US-10-324-184-14	Sequence 14, Appl
	15	17.2	13.0	22	16	US-10-179-826-3	Sequence 3, Appli
	16	17.2	13.0	24	16	US-10-179-826-4	Sequence 4, Appli
c	17	17.2	13.0	44	15	US-10-076-802-88	Sequence 88, Appl
c	18	17.2	13.0	44	16	US-10-367-169-88	Sequence 88, Appl
c	19	17	12.9	47	13	US-10-333-429-533	Sequence 533, App
c	20	17	12.9	47	16	US-10-349-143-735	Sequence 735, App
	21	17	12.9	48	10	US-09-993-346-228	Sequence 228, App
c	22	16.8	12.7	41	13	US-10-453-827-34	Sequence 34, Appl
c	23	16.8	12.7	50	10	US-09-993-346-551	Sequence 551, App
	24	16.6	12.6	31	10	US-09-877-478-4996	Sequence 4996, Ap
	25	16.6	12.6	31	13	US-10-342-902-4996	Sequence 4996, Ap
c	26	16.6	12.6	42	17	US-10-603-229-24	Sequence 24, Appl
	27	16.6	12.6	47	16	US-10-349-143-2420	Sequence 2420, Ap
c	28	16.6	12.6	50	10	US-09-993-346-400	Sequence 400, App
	29	16.4	12.4	31	17	US-10-258-728-7	Sequence 7, Appli
c	30	16.4	12.4	35	10	US-09-816-733-1	Sequence 1, Appli
c	31	16.4	12.4	35	13	US-10-440-067A-1	Sequence 1, Appli
c	32	16.4	12.4	41	12	US-10-035-833A-733	Sequence 733, App
c	33	16.4	12.4	41	12	US-10-035-833A-6324	Sequence 6324, Ap
c	34	16.4	12.4	41	15	US-10-292-418-36	Sequence 36, Appl
	35	16.4	12.4	44	9	US-09-874-547-2	Sequence 2, Appli
c	36	16.4	12.4	44	15	US-10-076-802-93	Sequence 93, Appl
c	37	16.4	12.4	44	16	US-10-367-169-93	Sequence 93, Appl
c	38	16.4	12.4	45	9	US-09-943-722-106	Sequence 106, App
c	39	16.4	12.4	46	9	US-09-795-006A-9	Sequence 9, Appli
	40	16.4	12.4	47	16	US-10-349-143-1555	Sequence 1555, Ap
c	41	16.4	12.4	47	16	US-10-349-143-3329	Sequence 3329, Ap
	42	16.4	12.4	50	9	US-09-978-295A-88	Sequence 88, Appl
	43	16.4	12.4	50	9	US-09-813-824A-20	Sequence 20, Appl
	44	16.4	12.4	50	9	US-09-978-697-88	Sequence 88, Appl
	45	16.4	12.4	50	9	US-09-978-192A-88	Sequence 88, Appl

ALIGNMENTS

RESULT 1
US-10-349-143-2729
; Sequence 2729, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CP1
; CURRENT APPLICATION NUMBER: US/10/349,143
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US/09/422,978
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2729
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-16284-389 : polymorphic base T or C
US-10-349-143-2729

Query Match 15.3%; Score 20.2; DB 16; Length 47;
Best Local Similarity 65.1%; Pred. No. 5.7e+03;
Matches 28; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 17 AGGACCTGCTGCACGACTGCGTGAGTTTGCAGCCCTGCATGCA 59

```
Db          5 AGGTCCTGCAGCAGTATGGYACCAGTTTGAAAGCCTCCATCCA 47


RESULT 2
US-10-097-044A-18/c
; Sequence 18, Application US/10097044A
; Publication No. US20030143220A1
;   GENERAL INFORMATION:
;        APPLICANT: Capon, Daniel J.
;                   Gregory, Timothy J.
;        TITLE OF INVENTION: Adheson Variants
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genentech, Inc.
;             STREET: 460 Point San Bruno Blvd
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: patin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/097,044A
;             FILING DATE: 28-May-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/457,918
;             FILING DATE: 1-JUN-1995
;             APPLICATION NUMBER: 08/236311
;             FILING DATE: 02-MAY-1994
;             APPLICATION NUMBER: 07/936190
;             FILING DATE: 26-AUG-1992
;             APPLICATION NUMBER: 07/842777
;             FILING DATE: 18-FEB-1992
;             APPLICATION NUMBER: 07/250785
;             FILING DATE: 28-SEP-1988
;             APPLICATION NUMBER: 07/104329
;             FILING DATE: 02-OCT-1987
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kubinec, Jeffrey S.
;             REGISTRATION NUMBER: 36,575
;             REFERENCE/DOCKET NUMBER: P0444P1C3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415/225-8228
;             TELEFAX: 415/952-9881
;             TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 18:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 47 bases
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-097-044A-18

  Query Match             14.5%;  Score 19.2;  DB 15;  Length 47;
  Best Local Similarity   67.5%;  Pred. No. 1.2e+04;
  Matches   27;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCC 113
              |||||| || | |||| |    |||||| |||| |   ||
Db         46 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCC 7


RESULT 3
US-10-097-044A-15
; Sequence 15, Application US/10097044A
; Publication No. US20030143220A1
;   GENERAL INFORMATION:
;        APPLICANT: Capon, Daniel J.
;                   Gregory, Timothy J.
;        TITLE OF INVENTION: Adheson Variants
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genentech, Inc.
;             STREET: 460 Point San Bruno Blvd
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: patin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/097,044A
;             FILING DATE: 28-May-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/457,918
;             FILING DATE: 1-JUN-1995
;             APPLICATION NUMBER: 08/236311
;             FILING DATE: 02-MAY-1994
;             APPLICATION NUMBER: 07/936190
;             FILING DATE: 26-AUG-1992
;             APPLICATION NUMBER: 07/842777
;             FILING DATE: 18-FEB-1992
;             APPLICATION NUMBER: 07/250785
;             FILING DATE: 28-SEP-1988
;             APPLICATION NUMBER: 07/104329
;             FILING DATE: 02-OCT-1987
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kubinec, Jeffrey S.
;             REGISTRATION NUMBER: 36,575
;             REFERENCE/DOCKET NUMBER: P0444P1C3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415/225-8228
;             TELEFAX: 415/952-9881
;             TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 15:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 47 bases
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-097-044A-15

  Query Match             14.4%;  Score 19;  DB 15;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 1.4e+04;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTG 108
              |||||| || | |||| |    |||||| |||| |
Db         12 CTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAG 46


RESULT 4
US-10-371-404-20
; Sequence 20, Application US/10371404
; Publication No. US20030194807A1
;   GENERAL INFORMATION:
;        APPLICANT: Crea, Roberto
;        TITLE OF INVENTION: Walk-Through Mutagenesis
;        NUMBER OF SEQUENCES: 59
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;             STREET: 2 Militia Drive
;             CITY: Lexington
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02173
```

```
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/371,404
;             FILING DATE: 20-Feb-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/453,623
;             FILING DATE: 30-May-1995
;             APPLICATION NUMBER: US 07/930,600
;             FILING DATE: 05-APR-1991
;             APPLICATION NUMBER: PCT/US91/02362
;             FILING DATE: 05-APR-1991
;             APPLICATION NUMBER: US 07/505,314
;             FILING DATE: 05-APR-1990
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brook, David E.
;             REGISTRATION NUMBER: 22,592
;             REFERENCE/DOCKET NUMBER: RC90-01AY
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 861-6240
;             TELEFAX: (617) 861-9540
;    INFORMATION FOR SEQ ID NO: 20:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 48 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: unknown
;        SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-371-404-20

  Query Match             13.8%;  Score 18.2;  DB 15;  Length 48;
  Best Local Similarity   48.9%;  Pred. No. 2.5e+04;
  Matches   23;  Conservative    8;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         64 TTCTGCGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTG 110
              | ||||||| | :||:||: :|| | |    |  ||: ::| : ||
Db          1 TACTGCGCGCGTARCTMCTMTRGCAGCASTTSGTMCTYCKMCKYTTG 47


RESULT 5
US-10-131-827-1070
; Sequence 1070, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1070
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-131-827-1070

  Query Match             13.6%;  Score 18;  DB 16;  Length 50;
  Best Local Similarity   70.6%;  Pred. No. 2.9e+04;
  Matches   24;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         51 CTGCATGCACACGTTCTGCGCGGCTTGCTACTCG 84
              |||  ||  ||   |||||| | ||||||| | |
Db         10 CTGTGTGAACTGTTTCTGCACTGCTTGCTAATGG 43


RESULT 6
US-10-076-802-81/c
; Sequence 81, Application US/10076802
; Publication No. US20030162249A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Jeff
;  APPLICANT:  Buechler, Joe
;  APPLICANT:  Veeramallu, Uday K
;  TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION
;  FILE REFERENCE: 014907-004700US
;  CURRENT APPLICATION NUMBER: US/10/076,802
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
;   OTHER INFORMATION: Heavy Chain Specific PCR Primer 188
US-10-076-802-81

  Query Match             13.5%;  Score 17.8;  DB 15;  Length 44;
  Best Local Similarity   75.9%;  Pred. No. 3.4e+04;
  Matches   22;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         20 ACCTGCTGCACGACTGCGTGAGTTTGCAG 48
              ||| ||||||| | || |||| |||||
Db         35 ACCAGCTGCACTTCCGCTTGAGCCTGCAG 7


RESULT 7
US-10-076-802-110/c
; Sequence 110, Application US/10076802
; Publication No. US20030162249A1
; GENERAL INFORMATION:
;  APPLICANT: Gray, Jeff
;  APPLICANT:  Buechler, Joe
;  APPLICANT:  Veeramallu, Uday K
;  TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION
;  FILE REFERENCE: 014907-004700US
;  CURRENT APPLICATION NUMBER: US/10/076,802
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 110
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
;   OTHER INFORMATION: Heavy Chain Specific PCR Primer 208
US-10-076-802-110

  Query Match             13.5%;  Score 17.8;  DB 15;  Length 44;
  Best Local Similarity   75.9%;  Pred. No. 3.4e+04;
  Matches   22;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         20 ACCTGCTGCACGACTGCGTGAGTTTGCAG 48
              ||| ||||||| | || |||| |||||
Db         35 ACCAGCTGCACTTCCGCTTGAGCCTGCAG 7


RESULT 8
US-10-076-802-111/c
; Sequence 111, Application US/10076802
; Publication No. US20030162249A1
```

; GENERAL INFORMATION:
; APPLICANT: Gray, Jeff
; APPLICANT: Buechler, Joe
; APPLICANT: Veeramallu, Uday K
; TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION
; FILE REFERENCE: 014907-004700US
; CURRENT APPLICATION NUMBER: US/10/076,802
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 44
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
; OTHER INFORMATION: Heavy Chain Specific PCR Primer 526
US-10-076-802-111

Query Match 13.5%; Score 17.8; DB 15; Length 44;
Best Local Similarity 75.9%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 20 ACCTGCTGCACGACTGCGTGAGTTTGCAG 48
||| ||||||| | || |||| |||||
Db 35 ACCACCTGCACGTCCGCTTGAGCCTGCAG 7

RESULT 9
US-10-367-169-81/c
; Sequence 81, Application US/10367169
; Publication No. US20030228660A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Jeff
; APPLICANT: Buechler, Joe
; APPLICANT: Veeramallu, Uday Kumar
; TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR POLYPEPTIDE EXPRESSION AND POLYPE
; TITLE OF INVENTION: DISPLAY LIBRARIES
; FILE REFERENCE: 11055US02
; CURRENT APPLICATION NUMBER: US/10/367,169
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 10/076,802
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 44
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
; OTHER INFORMATION: Heavy Chain Specific PCR Primer 188
US-10-367-169-81

Query Match 13.5%; Score 17.8; DB 16; Length 44;
Best Local Similarity 75.9%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 20 ACCTGCTGCACGACTGCGTGAGTTTGCAG 48
||| ||||||| | || |||| |||||
Db 35 ACCAGCTGCACTTCCGCTTGAGCCTGCAG 7

RESULT 10
US-10-367-169-110/c
; Sequence 110, Application US/10367169
; Publication No. US20030228660A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Jeff
; APPLICANT: Buechler, Joe
; APPLICANT: Veeramallu, Uday Kumar
; TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR POLYPEPTIDE EXPRESSION AND POLYPE
; TITLE OF INVENTION: DISPLAY LIBRARIES
; FILE REFERENCE: 11055US02
; CURRENT APPLICATION NUMBER: US/10/367,169
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 10/076,802
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 110
; LENGTH: 44
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
; OTHER INFORMATION: Heavy Chain Specific PCR Primer 208
US-10-367-169-110

Query Match 13.5%; Score 17.8; DB 16; Length 44;
Best Local Similarity 75.9%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 20 ACCTGCTGCACGACTGCGTGAGTTTGCAG 48
||| ||||||| | || |||| |||||
Db 35 ACCAGCTGCACTTCCGCTTGAGCCTGCAG 7

RESULT 11
US-10-367-169-111/c
; Sequence 111, Application US/10367169
; Publication No. US20030228660A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Jeff
; APPLICANT: Buechler, Joe
; APPLICANT: Veeramallu, Uday Kumar
; TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR POLYPEPTIDE EXPRESSION AND POLYPE
; TITLE OF INVENTION: DISPLAY LIBRARIES
; FILE REFERENCE: 11055US02
; CURRENT APPLICATION NUMBER: US/10/367,169
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 10/076,802
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 44
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
; OTHER INFORMATION: Heavy Chain Specific PCR Primer 526
US-10-367-169-111

Query Match 13.5%; Score 17.8; DB 16; Length 44;
Best Local Similarity 75.9%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 20 ACCTGCTGCACGACTGCGTGAGTTTGCAG 48
||| ||||||| | || |||| |||||
Db 35 ACCACCTGCACGTCCGCTTGAGCCTGCAG 7

RESULT 12
US-10-458-568-9/c
; Sequence 9, Application US/10458568
; Publication No. US20030204873A1
; GENERAL INFORMATION:
; APPLICANT: Schmidt, et al.
; TITLE OF INVENTION: Method For Producing Transgenic Plants With Modified 5-Aminolevul
; TITLE OF INVENTION: Acid Biosynthesis, And Method For Identifying 5-Aminolevulinic A
; TITLE OF INVENTION: Effectors
; FILE REFERENCE: 514413-3827
; CURRENT APPLICATION NUMBER: US/10/458,568
; CURRENT FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US/09/581,105

```
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: PCT/EP98/08028
;  PRIOR FILING DATE: 1998-12-10
;  PRIOR APPLICATION NUMBER: GERMANY 197 54 929.2
;  PRIOR FILING DATE: 1997-12-10
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: primer ALAS1
US-10-458-568-9

  Query Match           13.2%;  Score 17.4;  DB 13;  Length 32;
  Best Local Similarity  77.8%;  Pred. No. 4.4e+04;
  Matches   21;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         35 GCGTGAGTTTGCAGCCCTGCATGCACA 61
              | |  || ||| || ||||||||||||
Db         30 GTGCCAGATTGTAGTCCTGCATGCACA 4


RESULT 13
US-10-216-054A-14
; Sequence 14, Application US/10216054A
; Publication No. US20030144232A1
; GENERAL INFORMATION:
;  APPLICANT: CANCER RESEARCH VENTURES LIMITED
;  APPLICANT:  AGAMI, Reuven
;  APPLICANT:  BRUMMELKAMP, Thijn
;  TITLE OF INVENTION: EXPRESSION SYSTEM
;  FILE REFERENCE: KILBURN1100-1
;  CURRENT APPLICATION NUMBER: US/10/216,054A
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US 60/377,482
;  PRIOR FILING DATE: 2002-05-02
;  PRIOR APPLICATION NUMBER: UK 0130955.8
;  PRIOR FILING DATE: 2001-12-24
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 49
;   TYPE: RNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Predicted stem loop transcript
US-10-216-054A-14

  Query Match           13.2%;  Score 17.4;  DB 15;  Length 49;
  Best Local Similarity  57.1%;  Pred. No. 4.6e+04;
  Matches   20;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         83 CGGGCTGGATGGAGCGCTCGTCCCTGTGTCCTACC 117
              |||||  |  :  || | :|| ||| | |:||: ||
Db         12 CGGGCCGAUUCAAGAGAUCGGCCCGGAGUCCUGCC 46


RESULT 14
US-10-324-184-14
; Sequence 14, Application US/10324184
; Publication No. US20030144239A1
; GENERAL INFORMATION:
;  APPLICANT: CANCER RESEARCH VENTURES LIMITED
;  APPLICANT:  AGAMI, Reuven
;  APPLICANT:  BRUMMELKAMP, Thijn
;  TITLE OF INVENTION: EXPRESSION SYSTEM
;  FILE REFERENCE: KILBURN1100-2
;  CURRENT APPLICATION NUMBER: US/10/324,184
;  CURRENT FILING DATE:  2002-12-19
;  PRIOR APPLICATION NUMBER: US 10/216,054
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/377,482
;  PRIOR FILING DATE: 2002-05-02
;  PRIOR APPLICATION NUMBER: UK 0130955.8
;  PRIOR FILING DATE: 2001-12-24
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 49
;   TYPE: RNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Sequence of the predicted stem loop transcript generated from
;   OTHER INFORMATION: the pSUPER-CDC20 vector as depicted in Figure 4.
US-10-324-184-14

  Query Match           13.2%;  Score 17.4;  DB 15;  Length 49;
  Best Local Similarity  57.1%;  Pred. No. 4.6e+04;
  Matches   20;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         83 CGGGCTGGATGGAGCGCTCGTCCCTGTGTCCTACC 117
              |||||  |  :  || | :|| ||| | |:||: ||
Db         12 CGGGCCGAUUCAAGAGAUCGGCCCGGAGUCCUGCC 46


RESULT 15
US-10-179-826-3
; Sequence 3, Application US/10179826
; Publication No. US20030235827A1
; GENERAL INFORMATION:
;  APPLICANT: McKeown, Brian
;  TITLE OF INVENTION: Methods and Compositions for Monitoring Primer Extension and Polym
;  TITLE OF INVENTION:  Detection Reactions
;  FILE REFERENCE: 13164US
;  CURRENT APPLICATION NUMBER: US/10/179,826
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer - organism matches to Ovis aries
US-10-179-826-3

  Query Match           13.0%;  Score 17.2;  DB 16;  Length 22;
  Best Local Similarity  86.4%;  Pred. No. 5e+04;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          8 TCATCTGCCAGGACCTGCTGCA 29
              |||| || ||||| ||||||||
Db          1 TCATGTGGCAGGAGCTGCTGCA 22


Search completed: May 26, 2004, 10:22:40
Job time : 577.198 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:12:53 ; Search time 916.382 Seconds
(without alignments)
4301.492 Million cell updates/sec

Title: US-10-048-046-1_COPY_997_1128
Perfect score: 132
Sequence: 1 acatgcatcatctgccagga.....ctacctgccgctgtcccgtg 132

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	19.2	14.5	49	9	AI284473	AI284473 qu82a12.x
	2	18.8	14.2	42	28	AZ788565	AZ788565 2M0035F14
	3	18.6	14.1	50	9	AU105739	AU105739 AU105739
	4	18.6	14.1	50	9	AU105740	AU105740 AU105740
	5	18.6	14.1	50	9	AU105741	AU105741 AU105741
	6	18.2	13.8	49	9	AA886434	AA886434 oj52d09.s
c	7	18.2	13.8	50	9	AU105564	AU105564 AU105564
	8	17.8	13.5	50	28	AZ432474	AZ432474 1M0217I21
	9	17.6	13.3	49	9	AI941380	AI941380 sc12e11.y
c	10	17.4	13.2	44	29	BX121674	BX121674 Danio rer
	11	17.4	13.2	48	28	AZ485794	AZ485794 1M0313K24
	12	17.4	13.2	50	28	BH624322	BH624322 1007078A0
c	13	17.2	13.0	46	14	H45385	H45385 yn99c12.r1
c	14	17.2	13.0	49	29	CG805551	CG805551 1118060E0
	15	17.2	13.0	50	9	AU105249	AU105249 AU105249
c	16	17.2	13.0	50	9	AU107541	AU107541 AU107541
c	17	17.2	13.0	50	28	AZ432474	AZ432474 1M0217I21
	18	17	12.9	50	9	AA948106	AA948106 on51a04.s
c	19	17	12.9	50	9	AU103176	AU103176 AU103176
c	20	17	12.9	50	9	AU105454	AU105454 AU105454
c	21	17	12.9	50	9	AU105455	AU105455 AU105455
c	22	17	12.9	50	9	AU105456	AU105456 AU105456
c	23	16.8	12.7	50	9	AL788155	AL788155 AL788155
	24	16.8	12.7	50	9	AU102427	AU102427 AU102427
c	25	16.6	12.6	43	9	AI744377	AI744377 tr10c11.x
	26	16.6	12.6	47	28	AZ795054	AZ795054 2M0049A11
	27	16.6	12.6	49	9	AA813480	AA813480 ai67a09.s
	28	16.6	12.6	50	9	AU105247	AU105247 AU105247
	29	16.6	12.6	50	9	AU105248	AU105248 AU105248
	30	16.6	12.6	50	9	AU105258	AU105258 AU105258
c	31	16.6	12.6	50	29	CG868958	CG868958 AA0263 Sa
c	32	16.4	12.4	35	29	DR20D7S	AL739074 Danio rer
c	33	16.4	12.4	37	14	H46233	H46233 yo18e09.s1
	34	16.4	12.4	42	9	AA878619	AA878619 oj21d06.s
c	35	16.4	12.4	46	10	BE729845	BE729845 601564766
c	36	16.4	12.4	46	29	BX171949	BX171949 Danio rer
	37	16.4	12.4	49	28	BH852307	BH852307 SALK_0744
	38	16.4	12.4	49	28	BH852314	BH852314 SALK_0744
c	39	16.4	12.4	50	9	AU105060	AU105060 AU105060
c	40	16.2	12.3	42	29	DME545174	AJ545174 Drosophil
	41	16.2	12.3	43	12	BG707367	BG707367 602672770
	42	16.2	12.3	46	12	BG177798	BG177798 602315050
	43	16.2	12.3	48	28	AZ311191	AZ311191 1M0026D04
c	44	16.2	12.3	50	9	AU102902	AU102902 AU102902
c	45	16.2	12.3	50	9	AU105570	AU105570 AU105570

ALIGNMENTS

RESULT 1
AI284473/c
LOCUS AI284473 49 bp mRNA linear EST 23-NOV-1998
DEFINITION qu82a12.x1 NCI_CGAP_Eso2 Homo sapiens cDNA clone IMAGE:1978558 3' similar to SW:PRPB_HUMAN P02814 PROLINE-RICH PEPTIDE P-B. ;contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AI284473
VERSION AI284473.1 GI:3922706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

```
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1978558"
                     /tissue_type="squamous cell carcinoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Eso2"
                     /note="Organ: esophagus; Vector: pCMV-SPORT6; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  Average insert size 1.1 kb. Life Technologies
                     catalog #: 11502-010"
ORIGIN

  Query Match           14.5%;  Score 19.2;  DB 9;  Length 49;
  Best Local Similarity   62.5%;  Pred. No. 5e+05;
  Matches   30;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy         84 GGGCTGGATGGAGCGCTCGTCCCTGTGTCCTACCTGCCGCTGTCCCGT 131
              |||| ||  || |   || |||| | | ||||| | || |   ||| |
Db         48 GGGCGGGGGGGGGATTTCTTCCCCGCGCCCTACTTCCCCCCAGCCCTT 1


RESULT 2
AZ788565
LOCUS       AZ788565                 42 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0035F14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0035F14 R, genomic survey sequence.
ACCESSION   AZ788565
VERSION     AZ788565.1  GI:12928495
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0035  row: F  column: 14
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 42.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0035F14"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match           14.2%;  Score 18.8;  DB 28;  Length 42;
  Best Local Similarity   76.7%;  Pred. No. 5.9e+05;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         39 GAGTTTGCAGCCCTGCATGCACACGTTCTG 68
              | ||||| | ||||||||  ||||||  ||
Db          7 GTGTTTGAATCCCTGCATCTACACGTGATG 36


RESULT 3
AU105739
LOCUS       AU105739                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105739 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC01265, mRNA sequence.
ACCESSION   AU105739
VERSION     AU105739.1  GI:13555260
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC01265"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match           14.1%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   84.0%;  Pred. No. 7.1e+05;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        107 TGTGTCCTACCTGCCGCTGTCCCGT 131
```

```
                ||||||||  | ||||||||||| |
Db         10 TGTGTCCTGTCAGCCGCTGTCCCCT 34


RESULT 4
AU105740
LOCUS       AU105740                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105740 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI05409, mRNA sequence.
ACCESSION   AU105740
VERSION     AU105740.1  GI:13555261
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05409"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             14.1%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   84.0%;  Pred. No. 7.1e+05;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        107 TGTGTCCTACCTGCCGCTGTCCCGT 131
                ||||||||  | ||||||||||| |
Db         17 TGTGTCCTGTCAGCCGCTGTCCCCT 41


RESULT 5
AU105741
LOCUS       AU105741                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105741 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            W1659, mRNA sequence.
ACCESSION   AU105741
VERSION     AU105741.1  GI:13555262
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="W1659"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             14.1%;  Score 18.6;  DB 9;  Length 50;
  Best Local Similarity   84.0%;  Pred. No. 7.1e+05;
  Matches   21;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        107 TGTGTCCTACCTGCCGCTGTCCCGT 131
                ||||||||  | ||||||||||| |
Db         11 TGTGTCCTGTCAGCCGCTGTCCCCT 35


RESULT 6
AA886434
LOCUS       AA886434                  49 bp    mRNA    linear   EST 30-MAR-1998
DEFINITION  oj52d09.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1501937 3'
            similar to TR:Q04118 Q04118 SALIVARY PROLINE-RICH GLYCOPROTEIN G1
            PRECURSOR. ;contains MER22.t3 MER22 repetitive element ;, mRNA
            sequence.
ACCESSION   AA886434
VERSION     AA886434.1  GI:3001542
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1501937"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid3"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
```

I and Eco RI sites of the modified pT7T3 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 13.8%; Score 18.2; DB 9; Length 49;
Best Local Similarity 66.7%; Pred. No. 8.8e+05;
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy          83 CGGGCTGGATGGAGCGCTCGTCCCTGTGTCCTACCTGCC 121
               || | |||  || | || | || || | ||| ||| |||
Db           1 CGCGGTGGGGGGGGGGCCCCTCTCTCTCTCCAACCCGCC 39
```

RESULT 7
AU105564/c
LOCUS AU105564 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU105564 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone COL00322, mRNA sequence.
ACCESSION AU105564
VERSION AU105564.1 GI:13555085
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL00322"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 13.8%; Score 18.2; DB 9; Length 50;
Best Local Similarity 61.7%; Pred. No. 8.8e+05;
Matches 29; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy          84 GGGCTGGATGGAGCGCTCGTCCCTGTGTCCTACCTGCCGCTGTCCCG 130
               ||||||    ||  |  |  |  |||||| ||| || |||||| |
Db          48 GGGCTGCTGTGACTGTCAGCTCTACTGTCCTTCCTACCTCTGTCCAG 2
```

RESULT 8
AZ432474
LOCUS AZ432474 50 bp DNA linear GSS 03-OCT-2000
DEFINITION 1M0217I21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0217I21 R, genomic survey sequence.
ACCESSION AZ432474
VERSION AZ432474.1 GI:10556487
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 50)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0217 row: I column: 21
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 50.
FEATURES Location/Qualifiers
source 1. .50
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0217I21"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 13.5%; Score 17.8; DB 28; Length 50;
Best Local Similarity 75.9%; Pred. No. 1.1e+06;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          34 TGCGTGAGTTTGCAGCCCTGCATGCACAC 62
               |||| | || ||||  | ||||| |||||
Db           4 TGCGCGCGTGTGCATGCGTGCATACACAC 32
```

RESULT 9
AI941380
LOCUS AI941380 49 bp mRNA linear EST 30-NOV-2001
DEFINITION sc12e11.y1 Gm-c1013 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1013-453 5' similar to SW:CB12_PETHY P13869 CHLOROPHYLL A-B BINDING PROTEIN OF LHCI TYPE II PRECURSOR ;, mRNA sequence.
ACCESSION AI941380
VERSION AI941380.1 GI:5688365
KEYWORDS EST.
SOURCE Glycine max (soybean)

```
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand This clone is
            available through: ResGen, Invitrogen Corp. 2130 South Memorial
            Parkway Huntsville, AL 35801 For further information call:
            (800)-533-4363 or contact via email: ccu@resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1013-453"
                     /tissue_type="Whole seedlings, 2-3 week old seedlings,
                     greenhouse grown"
                     /lab_host="XL10-Gold"
                     /clone_lib="Gm-c1013"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from whole seedlings of 2-3 week old greenhouse grown
                     plants. The cDNA library was prepared using the Stratagene
                     pBluescript II XR cDNA library construction kit.
                     Complementary DNA was synthesized from mRNA using a primer
                     consisting of a poly (dT) sequence with a XhoI restriction
                     site. EcoRI adapters were ligated to the blunt-ended cDNA
                     fragments followed by XhoI digestion. The cDNA fragments
                     were directionally cloned into the EcoRI-XhoI restriction
                     site of the pBluescript vector. The ligated cDNA fragments
                     were transformed into XL10-Gold host cells. This library
                     was constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
ORIGIN

  Query Match             13.3%;  Score 17.6;  DB 9;  Length 49;
  Best Local Similarity   65.0%;  Pred. No. 1.2e+06;
  Matches   26;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         40 AGTTTGCAGCCCTGCATGCACACGTTCTGCGCGGCTTGCT 79
              | | ||   |||||    || ||| | ||| | |||||| |
Db         10 ATTGTGGTTCCCTGGCAGCTCACCTCCTGAGTGGCTTGAT 49


RESULT 10
BX121674/c
LOCUS       BX121674                 44 bp    DNA     linear   GSS 28-JAN-2003
DEFINITION  Danio rerio genomic clone DKEY-61M14, genomic survey sequence.
ACCESSION   BX121674
VERSION     BX121674.1  GI:27952599
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 61M14. 61M14 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-61M14"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

  Query Match             13.2%;  Score 17.4;  DB 29;  Length 44;
  Best Local Similarity   61.8%;  Pred. No. 1.3e+06;
  Matches   21;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         76 TGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGT 109
              ||   |||||   || ||  | |  |||||||
Db         41 TGTCGCTCGGNNNGGTTGNNGTGNNCGTCCCTCT 8


RESULT 11
AZ485794
LOCUS       AZ485794                 48 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0313K24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0313K24 F, genomic survey sequence.
ACCESSION   AZ485794
VERSION     AZ485794.1  GI:10651939
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0313  row: K  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0313K24"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
```

/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 13.2%; Score 17.4; DB 28; Length 48;
Best Local Similarity 62.8%; Pred. No. 1.4e+06;
Matches 27; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy         49 CCCTGCATGCACACGTTCTGCGCGGCTTGCTACTCGGGCTGGA 91
              ||||| ||| || | || ||| ||| ||     | |  ||||
Db          2 CCCTGGGTGCCCATGATCGGCGGGGCCTGGCCTGCAGCTTGGA 44
```

RESULT 12
BH624322
LOCUS BH624322 50 bp DNA linear GSS 30-JAN-2002
DEFINITION 1007078A02.2EL_y2 1007 - RescueMu Grid H Zea mays genomic, genomic survey sequence.
ACCESSION BH624322
VERSION BH624322.1 GI:18437573
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 50)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007078 row: 30
Class: transposon-tagged.
FEATURES Location/Qualifiers
source 1. .50
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1007 - RescueMu Grid H"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 13.2%; Score 17.4; DB 28; Length 50;
Best Local Similarity 62.8%; Pred. No. 1.4e+06;
Matches 27; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy         89 GGATGGAGCGCTCGTCCCTGTGTCCTACCTGCCGCTGTCCCGT 131
              || |||||| |||||| | || ||    || ||   ||  ||
Db          1 GGCTGGAGCAGTCGTCCTTTTGGCCGGATTGTCGTCTTCTTGT 43
```

RESULT 13
H45385/c
LOCUS H45385 46 bp mRNA linear EST 31-JUL-1995
DEFINITION yn99c12.r1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone IMAGE:176566 5' similar to SP:TFS2_MOUSE P10712 TRANSCRIPTION FACTOR S-II ;, mRNA sequence.
ACCESSION H45385
VERSION H45385.1 GI:921437
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 46)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .46
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3838762"
/db_xref="taxon:9606"
/clone="IMAGE:176566"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares adult brain N2b5HB55Y"
/note="Organ: brain; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector

(Pharmacia). Library went through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M.Fatima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."

ORIGIN

Query Match 13.0%; Score 17.2; DB 14; Length 46;
Best Local Similarity 65.8%; Pred. No. 1.5e+06;
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy          5 GCATCATCTGCCAGGACCTGCTGCACGACTGCGTGAGT 42
              | | | ||| ||||||| || ||   ||||  |  |||
Db         44 GGAGCTTCTTCCAGGACTTGATGAGAGACTTGGCCAGT 7
```

RESULT 14
CG805551/c
LOCUS CG805551 49 bp DNA linear GSS 10-NOV-2003
DEFINITION 1118060E02.2EL_y1 1118 - RescueMu Grid S Zea mays genomic, genomic survey sequence.
ACCESSION CG805551
VERSION CG805551.1 GI:38241970
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 49)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1118060 row: 36
Class: transposon-tagged.
FEATURES Location/Qualifiers
source 1. .49
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1118 - RescueMu Grid S"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 13.0%; Score 17.2; DB 29; Length 49;
Best Local Similarity 73.3%; Pred. No. 1.5e+06;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy         23 TGCTGCACGACTGCGTGAGTTTGCAGCCCT 52
              ||| ||| || || ||||  | ||||| ||
Db         34 TGCAGCATGAGTGTGTGATATGGCAGCACT 5
```

RESULT 15
AU105249
LOCUS AU105249 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU105249 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC05540, mRNA sequence.
ACCESSION AU105249
VERSION AU105249.1 GI:13554770
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC05540"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 13.0%; Score 17.2; DB 9; Length 50;
Best Local Similarity 60.9%; Pred. No. 1.6e+06;
Matches 28; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy         40 AGTTTGCAGCCCTGCATGCACACGTTCTGCGCGGCTTGCTACTCGG 85
              ||||| | |||| ||  | ||   | ||||||     ||| | | |
Db          1 AGTTTTCCGCCCGGCGCGGACGGCTGCTGCGCCCGCGGCTCCCCAG 46
```

Search completed: May 26, 2004, 09:17:48
Job time : 919.382 secs